

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using SW model

Run on: February 9, 2004, 16:14:15 ; Search time 22 Seconds
(without alignments)
2413.642 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815
Sequence: 1 MELALACRWGLLLALPPGA.....TFKGTPTANPEYLGLDPV 1255

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	US-08-625-101-2	Sequence 2, Appl
2	6815	100.0	1255	US-08-356-786-2	Sequence 2, Appl
3	6806	99.9	1255	US-08-467-083-68	Sequence 68, Appl
4	6806	99.9	1255	US-08-414-417B-68	Sequence 68, Appl
5	6806	99.9	1255	US-08-486-348A-68	Sequence 68, Appl
6	6806	99.9	1255	US-08-468-545B-68	Sequence 68, Appl
7	6806	99.9	1255	US-08-466-680B-68	Sequence 68, Appl
8	6806	99.9	1255	US-09-527-487-2	Sequence 2, Appl
9	6732	98.8	1255	US-08-484-438-8	Sequence 8, Appl
10	3630	53.3	782	US-09-146-283-4	Sequence 8, Appl
11	3630	53.3	782	US-08-579-823A-4	Sequence 4, Appl
12	3630	53.3	782	US-09-344-195-4	Sequence 4, Appl
13	3473	51.0	624	US-08-422-108-1	Sequence 1, Appl
14	3473	51.0	624	US-08-475-035-4	Sequence 4, Appl
15	3169	46.5	1210	US-08-484-438-7	Sequence 7, Appl
16	3169	46.5	1210	US-08-484-438-7	Sequence 7, Appl
17	3091	45.4	580	US-08-486-348A-69	Sequence 69, Appl
18	3091	45.4	580	US-08-466-680B-69	Sequence 69, Appl
19	3091	45.4	580	US-08-468-545B-69	Sequence 69, Appl
20	3091	45.4	580	US-08-484-438-2	Sequence 2, Appl
21	2871.5	42.1	1058	US-08-484-438-4	Sequence 4, Appl
22	2871.5	42.1	1058	US-07-978-895-4	Sequence 4, Appl
23	2441.5	35.8	1342	US-08-484-438-9	Sequence 9, Appl
24	2441.5	35.8	1342	US-08-473-119-4	Sequence 4, Appl
25	2441.5	35.8	1342	US-08-475-352-4	Sequence 4, Appl
26	2441.5	35.8	1342	US-08-475-352-4	Sequence 4, Appl
27	2434	35.7	1343	5183884-4	Patent No. 5183884

28	1878	27.6	419	4	US-09-630-155-2	Sequence 2, Appl
29	1646.5	24.2	705	2	US-08-456-647B-4	Sequence 4, Appl
30	1646.5	24.2	705	2	US-08-237-401A-4	Sequence 4, Appl
31	1532.5	22.5	644	1	US-08-336-708A-9	Sequence 9, Appl
32	1451	21.3	911	2	US-08-484-438-10	Sequence 10, Appl
33	1389.5	20.4	265	2	US-07-857-224B-66	Sequence 66, Appl
34	1235	18.1	279	2	US-08-701-191A-14	Sequence 14, Appl
35	1188	17.4	541	2	US-08-484-438-6	Sequence 6, Appl
36	1165.5	17.1	265	2	US-07-857-224B-65	Sequence 65, Appl
37	1028.5	15.1	478	4	US-09-570-454-2	Sequence 2, Appl
38	1028.5	15.1	478	4	US-09-867-521-2	Sequence 2, Appl
39	820	12.0	264	2	US-07-857-224B-67	Sequence 67, Appl
40	715	10.5	1382	4	US-08-737-715-2	Sequence 2, Appl
41	715	10.5	1382	4	US-09-457-040B-7	Sequence 7, Appl
42	653	9.6	1367	2	US-08-625-819-2	Sequence 2, Appl
43	651	9.6	1367	2	US-08-249-687C-2	Sequence 2, Appl
44	651	9.6	1367	3	US-08-746-559A-2	Sequence 2, Appl
45	641	9.4	1367	4	US-08-864-641B-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-625-101-2
Sequence 2, Application US/08625101
Patent No. 5868445
GENERAL INFORMATION:
APPLICANT: Cnoveer, Martin A.
APPLICANT: Disib, Mary L.
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/625,101
FILING DATE: 01-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-625-101-2
Query Match 100.0%; Score 6815; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MELALACRWGLLLALPPGAATGCTGTDKRLPASPETHLDMRLHYOCQOVQGNL 60
Db 1 MELALACRWGLLLALPPGAATGCTGTDKRLPASPETHLDMRLHYOCQOVQGNL 60

```

QY 61 ELTYLPTNASTSLFLODIQEVQGYVLI.AHNQVROVPLQRLRI.VRGTOU.FEDNYALAVLDNG 120
DB 61 ELTYLPTNASTSLFLODIQEVQGYVLI.AHNQVROVPLQRLRI.VRGTOU.FEDNYALAVLDNG 120
QY 121 DELNNTTPVTGASPGGIRELOLSLFEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DELNNTTPVTGASPGGIRELOLSLFEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180
QY 121 DELNNTTPVTGASPGGIRELOLSLFEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DELNNTTPVTGASPGGIRELOLSLFEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDTNRSPACSPSCPMCKSGRCWGBSSSDCOSLFTTVAGAGCARCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSPACSPSCPMCKSGRCWGBSSSDCOSLFTTVAGAGCARCKGPLPTDCCHQC 240
QY 241 AACCTGPKSHDCLAHLFHNHSGICEHLCPALVTYNTDTPESMNPBGRYTFGASCVTACP 300
DB 241 AACCTGPKSHDCLAHLFHNHSGICEHLCPALVTYNTDTPESMNPBGRYTFGASCVTACP 300
QY 301 YNLTSTDVGSCTIVCPAHNEVTAEDTORCEKSKPCAVCYGLGMEHLREVAATVSAN 360
DB 301 YNLTSTDVGSCTIVCPAHNEVTAEDTORCEKSKPCAVCYGLGMEHLREVAATVSAN 360
QY 361 IOEPACCKIKFGLAFPLPESFDGPASNTAPLOPBOLOVEETLEITGYLIISAMPDLP 420
DB 361 IOEPACCKIKFGLAFPLPESFDGPASNTAPLOPBOLOVEETLEITGYLIISAMPDLP 420
QY 421 DLSVFNOLQYIRGRILHNHNGVSLTLQGLISWLGSLRSLRSGSLALIHNTHLCPVHTV 480
DB 421 DLSVFNOLQYIRGRILHNHNGVSLTLQGLISWLGSLRSLRSGSLALIHNTHLCPVHTV 480
QY 481 PMOOLFRNHPQALHTNRPEDDEVGBGLACHOLCARGHCWGPPTQVNCOSOTLRQEC 540
DB 481 PMOOLFRNHPQALHTNRPEDDEVGBGLACHOLCARGHCWGPPTQVNCOSOTLRQEC 540
QY 541 VEECRVLOGLPREVYNAHCLPCHPECOPONGSVTCGFPADOCVACAHYKDPFCVARC 600
DB 541 VEECRVLOGLPREVYNAHCLPCHPECOPONGSVTCGFPADOCVACAHYKDPFCVARC 600
QY 601 PSQVKDPLSTWPIWKPEDEBGAQCPINCTHSCVDLDDKCPABQASPLTISAVG 660
DB 601 PSQVKDPLSTWPIWKPEDEBGAQCPINCTHSCVDLDDKCPABQASPLTISAVG 660
QY 661 ILLVVLGVVPGILIKRROOKIRYTMRRLLQETELVEPLTPSGAMNOMRLKETEL 720
DB 661 ILLVVLGVVPGILIKRROOKIRYTMRRLLQETELVEPLTPSGAMNOMRLKETEL 720
QY 721 RYKVLGSGAFGYVYKGIWI.PDGENVKI.PVALIKVRENTSPKANKELIDRAVYMAVGSF 780
DB 721 RYKVLGSGAFGYVYKGIWI.PDGENVKI.PVALIKVRENTSPKANKELIDRAVYMAVGSF 780
QY 781 YVSRLLGICLTSFVQVLTQMPYGCLL.DHVRENRGLSGODLNMCMQIAKMSYLEDVR 840
DB 781 YVSRLLGICLTSFVQVLTQMPYGCLL.DHVRENRGLSGODLNMCMQIAKMSYLEDVR 840
QY 841 LVHRDLAARVLYKSPHNHVKITDREGARLLDIDETETHAOGKPIPMMALESTLRERFT 900
DB 841 LVHRDLAARVLYKSPHNHVKITDREGARLLDIDETETHAOGKPIPMMALESTLRERFT 900
QY 901 HOSDWSYGVTWELMTFGAKPYDGI.PAREIPDLLEKGERLPOPPICITIDVYMMVMCM 960
DB 901 HOSDWSYGVTWELMTFGAKPYDGI.PAREIPDLLEKGERLPOPPICITIDVYMMVMCM 960
QY 961 IDSECRPRELVESEFARMADPORFVITONEDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRELVESEFARMADPORFVITONEDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 BEYLVPOOGFPCPDPAQAGMVAHRRSSSTRSGGGLTLGLEPSEBEAPRSGLASBG 1080
DB 1021 BEYLVPOOGFPCPDPAQAGMVAHRRSSSTRSGGGLTLGLEPSEBEAPRSGLASBG 1080
QY 1081 AGSDVFPDGLGMAAKGLOS.LPTHDPSPLOYSBPTVPLPSERTDGYVAPLTCSPOPEYV 1140
DB 1081 AGSDVFPDGLGMAAKGLOS.LPTHDPSPLOYSBPTVPLPSERTDGYVAPLTCSPOPEYV 1140

```

```

QY 1141 NQDVAREPPSPREGELPAPAPAGATLERPKTISBKNGVVKDVAFPGAVENPEVLTPO 1200
DB 1141 NQDVAREPPSPREGELPAPAPAGATLERPKTISBKNGVVKDVAFPGAVENPEVLTPO 1200
QY 1201 GGAAPQHPHPAFSPAFNDLYYWDODPPRGAAPPSTFKGPTAEANEYGLDVPV 1255
DB 1201 GGAAPQHPHPAFSPAFNDLYYWDODPPRGAAPPSTFKGPTAEANEYGLDVPV 1255

RESULT 2
US-08-356-786-2
; Sequence 2, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; NUMBER OF SEQUENCES: 16
; ADDRESS: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-786-2

Query Match 100.0%; Score 6815; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db      181 LTLIDTNRSRACHPCSPMKSGSRCSGSSSDCOSLTRTVCAAGCARCKPPLPTDCHEOC 240
Qy      241 AAGCTGPRKSDCLACLFHNSGICELHCPALVYNTDTFESMNPREGRYTFGASCVTACP 300
Db      241 AAGCTGPRKSDCLACLFHNSGICELHCPALVYNTDTFESMNPREGRYTFGASCVTACP 300
Qy      301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSCPKCARVCGIGMEHLREVAVTSAN 360
Db      301 YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSCPKCARVCGIGMEHLREVAVTSAN 360
Qy      361 IOEPACCKIFGSLAFPLPESPDGPASNTAPLOPEOLVPELLEITGLYISAMPDLP 420
Db      361 IOEPACCKIFGSLAFPLPESPDGPASNTAPLOPEOLVPELLEITGLYISAMPDLP 420
Qy      421 DLSVFNQIVIRGRILHNGAVSLTQGLGISWLGRLSRLBSGLALIHNNHLCFVHTV 480
Db      421 DLSVFNQIVIRGRILHNGAVSLTQGLGISWLGRLSRLBSGLALIHNNHLCFVHTV 480
Qy      481 PMDOLFRPHQALLHTANRPEBECYGBGLACHQLCARGHGCPGPTQCNCSQPLRGQEC 540
Db      481 PMDOLFRPHQALLHTANRPEBECYGBGLACHQLCARGHGCPGPTQCNCSQPLRGQEC 540
Qy      541 VEEGRVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCAVCAHYDPPFCVARC 600
Db      541 VEEGRVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCAVCAHYDPPFCVARC 600
Qy      601 PSQVCPDLSYMPIMKFPDEBEGACOPCPINCTHSCVDLDKGCFAQRASPLTISIAYVG 660
Db      601 PSQVCPDLSYMPIMKFPDEBEGACOPCPINCTHSCVDLDKGCFAQRASPLTISIAYVG 660
Qy      661 ILVVVVLGVVGIILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
Db      661 ILVVVVLGVVGIILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
Qy      721 RKVAVLSGAGATYKGIWIPDGRNVKIPVAIKYLRNTSPKANEILDEAYVAVGASP 780
Db      721 RKVAVLSGAGATYKGIWIPDGRNVKIPVAIKYLRNTSPKANEILDEAYVAVGASP 780
Qy      781 YVSRLLGICLSTVQLVTLMPYGCILDHVENRGLSGODLWMCQIAGMSYLEDVR 840
Db      781 YVSRLLGICLSTVQLVTLMPYGCILDHVENRGLSGODLWMCQIAGMSYLEDVR 840
Qy      841 LVHRDLAARNVLVSPNHVKITDFGLARLLDIDETRYADGKVPKIMMALESLIRRET 900
Db      841 LVHRDLAARNVLVSPNHVKITDFGLARLLDIDETRYADGKVPKIMMALESLIRRET 900
Qy      901 HOSDVMYGVTVWELMTFGAPYDGIIPAREIPDLLEKGERLPPPICTIDVYIMVCKMM 960
Db      901 HOSDVMYGVTVWELMTFGAPYDGIIPAREIPDLLEKGERLPPPICTIDVYIMVCKMM 960
Qy      961 IDESCRPRFELVSEFRMARDPORFVYIOWEDIGPASPILSTFYSRLLEDMDGDLVDA 1020
Db      961 IDESCRPRFELVSEFRMARDPORFVYIOWEDIGPASPILSTFYSRLLEDMDGDLVDA 1020
Qy      1021 EBYLVPQGGFFCPDPAAGAGVHRRHRSSTRSGGDLTLGLEPSEBEAPRSLASEG 1080
Db      1021 EBYLVPQGGFFCPDPAAGAGVHRRHRSSTRSGGDLTLGLEPSEBEAPRSLASEG 1080
Qy      1081 AGSDVPRGDLGMAKGLQSLPTHDPSLQRYSEDPTVPLPSETDGYVAPLTCSPQREYV 1140
Db      1081 AGSDVPRGDLGMAKGLQSLPTHDPSLQRYSEDPTVPLPSETDGYVAPLTCSPQREYV 1140
Qy      1141 NQPDVPRPSPREGPLPAARPAATLERPKTSLPGKNGVYKDFAFAGGAVENREYLTPO 1200
Db      1141 NQPDVPRPSPREGPLPAARPAATLERPKTSLPGKNGVYKDFAFAGGAVENREYLTPO 1200
Qy      1201 GGAAPRHPRPAPSPADNTLYMDQDPRERGAPOSTFKGTPTAENPEYLGIDVAV 1255
Db      1201 GGAAPRHPRPAPSPADNTLYMDQDPRERGAPOSTFKGTPTAENPEYLGIDVAV 1255

```

RESULT 3
US-08-467-083-68

```

; Sequence 68, Application US/08467083
; Patent No. 5726023
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disib, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,083
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/414,417
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836 SEDANBERY
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-467-083-68

Query Match      99.9%; Score 6806; DB 1; Length 1255;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  MELNALCRWGLLALALPPGAASVQCTGTMKRLPASPEHLMRLHLYOGCOVVQGNL 60
Db      1  MELNALCRWGLLALALPPGAASVQCTGTMKRLPASPEHLMRLHLYOGCOVVQGNL 60
Qy      61  ELTYLPTNASLSFLQDIQEVQGYVLIHNOVROVPLQRLRIVRGTOJFEDNYALAVLDNG 120
Db      61  ELTYLPTNASLSFLQDIQEVQGYVLIHNOVROVPLQRLRIVRGTOJFEDNYALAVLDNG 120
Qy      121  DPLNNTPTVYGAPPGGRLRELOLSLTILKGVLIQNNPOLCYDDITLWKDIFHKNOOLA 180
Db      121  DPLNNTPTVYGAPPGGRLRELOLSLTILKGVLIQNNPOLCYDDITLWKDIFHKNOOLA 180
Qy      181  LTLIDTNRSRACHPCSPMKSGSRCSGSSSDCOSLTRTVCAAGCARCKPPLPTDCHEOC 240
Db      181  LTLIDTNRSRACHPCSPMKSGSRCSGSSSDCOSLTRTVCAAGCARCKPPLPTDCHEOC 240
Qy      241  AAGCTGPRKSDCLACLFHNSGICELHCPALVYNTDTFESMNPREGRYTFGASCVTACP 300
Db      241  AAGCTGPRKSDCLACLFHNSGICELHCPALVYNTDTFESMNPREGRYTFGASCVTACP 300
Qy      301  YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSCPKCARVCGIGMEHLREVAVTSAN 360
Db      301  YNYLSTDVGSCTLVCPPLHNOEYTAEDGTORCEKSCPKCARVCGIGMEHLREVAVTSAN 360
Qy      361  IOEPACCKIFGSLAFPLPESPDGPASNTAPLOPEOLVPELLEITGLYISAMPDLP 420

```

Db 361 IOEPACCKKI FGSGLAFPSFDGDPASNTAPLOPBOLOVFEITLBTGYLYISAMPDSLP 420
 Qy 421 DLSVFONLOVIRIRIANGAYSITLLOGIGISWGLASRLBSGLALIHNTLCEVHTV 480
 Db 421 DLSVFONLOVIRIRIANGAYSITLLOGIGISWGLASRLBSGLALIHNTLCEVHTV 480
 Qy 481 PWDOLFPRNPHOALLHTANRPEDECVGEGLACHOLCARHCWGPGPTQCVNCSQFLRGQEC 540
 Db 481 PWDOLFPRNPHOALLHTANRPEDECVGEGLACHOLCARHCWGPGPTQCVNCSQFLRGQEC 540
 Qy 541 VEECRVLOGLPREYVNAHCLPCHPCOPONGSVTCFPGPADOCVACAHYKDPFCVARC 600
 Db 541 VEECRVLOGLPREYVNAHCLPCHPCOPONGSVTCFPGPADOCVACAHYKDPFCVARC 600
 Qy 601 PGVVRKDLSTYMPIMKPPDEEGACQPCPINCSTHSCVDLDKGCAPABASPLTISIAYVG 660
 Db 601 PGVVRKDLSTYMPIMKPPDEEGACQPCPINCSTHSCVDLDKGCAPABASPLTISIAYVG 660
 Qy 661 ILLVVVLGVVFGILLKRRQOKIRKTYTRRLLOETELVBPITPSGAMPNOQMRILKETEL 720
 Db 661 ILLVVVLGVVFGILLKRRQOKIRKTYTRRLLOETELVBPITPSGAMPNOQMRILKETEL 720
 Qy 721 RRVKVLGSGAFGVYKGIWIPDEENKIPVALKYLENTSPKANKIILDEAYVMAGVSP 780
 Db 721 RRVKVLGSGAFGVYKGIWIPDEENKIPVALKYLENTSPKANKIILDEAYVMAGVSP 780
 Qy 781 YSRLLIGICLTSTVQOLVTOIMPYGLLDHYBNRGLSGODLLNMCQIAGMSYLEDVR 840
 Db 781 YSRLLIGICLTSTVQOLVTOIMPYGLLDHYBNRGLSGODLLNMCQIAGMSYLEDVR 840
 Qy 841 LVHRDLAARNVLYKSPBNHYKITDPLGLARLLIDETRYHADGKVPIMKMALESILRRPT 900
 Db 841 LVHRDLAARNVLYKSPBNHYKITDPLGLARLLIDETRYHADGKVPIMKMALESILRRPT 900
 Qy 901 HOSDWSYGVYTWELMTFGAKPYDGI PARI PDLLEKGERLPOPPCTIDVYIMKCMW 960
 Db 901 HOSDWSYGVYTWELMTFGAKPYDGI PARI PDLLEKGERLPOPPCTIDVYIMKCMW 960
 Qy 961 IOSECRPRELVESESRMARDPORFVIONEDLGPASPLDSTFYSLIEDDMGDLVDA 1020
 Db 961 IOSECRPRELVESESRMARDPORFVIONEDLGPASPLDSTFYSLIEDDMGDLVDA 1020
 Qy 1021 BEYLVOQGFPCDPAFAGAGVHHRSSSTRSGGGLTLGLSPSEEARSPPLAPSEG 1080
 Db 1021 BEYLVOQGFPCDPAFAGAGVHHRSSSTRSGGGLTLGLSPSEEARSPPLAPSEG 1080
 Qy 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLORYSEDPVPLPSETDGVNAELTCSPOPEYV 1140
 Db 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLORYSEDPVPLPSETDGVNAELTCSPOPEYV 1140
 Qy 1141 NOPDVAPPPRPPRGLPARPAGATLERPKTSPGKNGVYKDVAFAGAVENPEYLTPO 1200
 Db 1141 NOPDVAPPPRPPRGLPARPAGATLERPKTSPGKNGVYKDVAFAGAVENPEYLTPO 1200
 Qy 1201 GGAAPRPHRPPASPAFNDLYWDDPBERGAPSTFKGTPLAENPEYGLDVPV 1255
 Db 1201 GGAAPRPHRPPASPAFNDLYWDDPBERGAPSTFKGTPLAENPEYGLDVPV 1255

RESULT 4
 US-08-414-17B-68
 Sequence 68, Application US/0841417B
 Patent No. 5801005
 GENERAL INFORMATION:
 APPLICANT: Cheever, Martin A.
 APPLICANT: Dils, Mary L.
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
 TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
 NUMBER OF SEQUENCES: 69
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Seed and Berry LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle
 STATE: Washington
 COUNTRY: US
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/414,417B
 FILING DATE: 31-MAR-1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Sharkey, Richard G.
 REGISTRATION NUMBER: 32,629
 REFERENCE/DOCKET NUMBER: 920010.448C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 68:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1255 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-414-17B-68

Query Match 99.9%; Score 6806; DB 1; Length 1255;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 METALCEWGLLALLPFGAASVQVCTGDMKRLRSPSETHIDMLRHLVYOGCQVVOGNL 60
 Db 1 METALCEWGLLALLPFGAASVQVCTGDMKRLRSPSETHIDMLRHLVYOGCQVVOGNL 60
 Qy 61 ELTYLPTNASTSFLDIOGVGVYLAHQVQVPLQRLRIYRGTOLEFENVYALAVDNG 120
 Db 61 ELTYLPTNASTSFLDIOGVGVYLAHQVQVPLQRLRIYRGTOLEFENVYALAVDNG 120
 Qy 121 DEANNTPTVAGSPGGLRELOLSLTELKGVLIQNPOLCYODITLWMDIFHKNNOLA 180
 Db 121 DEANNTPTVAGSPGGLRELOLSLTELKGVLIQNPOLCYODITLWMDIFHKNNOLA 180
 Qy 181 LTLIDTNRACHPCSPCKSGRCWSESSDCSLTRTYCAGGCANCKGFLPDDCHEOC 240
 Db 181 LTLIDTNRACHPCSPCKSGRCWSESSDCSLTRTYCAGGCANCKGFLPDDCHEOC 240
 Qy 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTESMNPBGRVTFGASCVTACP 300
 Db 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTESMNPBGRVTFGASCVTACP 300
 Qy 301 YNLTSDVGSCTLVCPHNOEYTAEDGTORCEKSPCARVCYGLGMEHLREYAVTSAN 360
 Db 301 YNLTSDVGSCTLVCPHNOEYTAEDGTORCEKSPCARVCYGLGMEHLREYAVTSAN 360
 Qy 361 IOEPACCKKI FGSGLAFPSFDGDPASNTAPLOPBOLOVFEITLBTGYLYISAMPDSLP 420
 Db 361 IOEPACCKKI FGSGLAFPSFDGDPASNTAPLOPBOLOVFEITLBTGYLYISAMPDSLP 420
 Qy 421 DLSVFONLOVIRIRIANGAYSITLLOGIGISWGLASRLBSGLALIHNTLCEVHTV 480
 Db 421 DLSVFONLOVIRIRIANGAYSITLLOGIGISWGLASRLBSGLALIHNTLCEVHTV 480
 Qy 481 PWDOLFPRNPHOALLHTANRPEDECVGEGLACHOLCARHCWGPGPTQCVNCSQFLRGQEC 540
 Db 481 PWDOLFPRNPHOALLHTANRPEDECVGEGLACHOLCARHCWGPGPTQCVNCSQFLRGQEC 540
 Qy 541 VEECRVLOGLPREYVNAHCLPCHPCOPONGSVTCFPGPADOCVACAHYKDPFCVARC 600
 Db 541 VEECRVLOGLPREYVNAHCLPCHPCOPONGSVTCFPGPADOCVACAHYKDPFCVARC 600
 Qy 601 PGVVRKDLSTYMPIMKPPDEEGACQPCPINCSTHSCVDLDKGCAPABASPLTISIAYVG 660


```

Db 601 PSQVPLSYMPKPEDEBGAQCPICNCTHSVDLDKCGCPAEQAPLTSISAVVG 660
Qy 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOAMRIKETEL 720
Db 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOAMRIKETEL 720
Qy 721 RKVKVLGSGAFGTYYKGIWIPDGENVKIPVAIKYLRRENTSPKANKIIDEAYVAVGSP 780
Db 721 RKVKVLGSGAFGTYYKGIWIPDGENVKIPVAIKYLRRENTSPKANKIIDEAYVAVGSP 780
Qy 781 YVSRLLGICLTSTVOLVTOAMPYGCILDHVRENRGLSGODLNNWMOIAKMSYLEDR 840
Db 781 YVSRLLGICLTSTVOLVTOAMPYGCILDHVRENRGLSGODLNNWMOIAKMSYLEDR 840
Qy 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETRYHADGSKVPIKMMALSIILRRFT 900
Db 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETRYHADGSKVPIKMMALSIILRRFT 900
Qy 901 HOSDVMSYGVYVWELMTFGAKPYDGI PAKEIPDLLEKGERLPQPICTIDVYMIWKCMM 960
Db 901 HOSDVMSYGVYVWELMTFGAKPYDGI PAKEIPDLLEKGERLPQPICTIDVYMIWKCMM 960
Qy 961 IDSECRPFRELVSEFSRMAARDPQRFVYVIONEDLPASPLDSTFYRSILLEDMDGLDVA 1020
Db 961 IDSECRPFRELVSEFSRMAARDPQRFVYVIONEDLPASPLDSTFYRSILLEDMDGLDVA 1020
Qy 1021 EEYLVPQGFPCPDPAFGAGVWHRHRSSSTRSGGGLTLGLBSEBEEAPRSPAPSEB 1080
Db 1021 EEYLVPQGFPCPDPAFGAGVWHRHRSSSTRSGGGLTLGLBSEBEEAPRSPAPSEB 1080
Qy 1081 AGSDVFPDDLDGMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEY 1140
Db 1081 AGSDVFPDDLDGMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEY 1140
Qy 1141 NOPDVRPQPPREBEPPLPAARPAAGATLRRPKTLSPGKGVVQDVAPAGAVENPEYLPQ 1200
Db 1141 NOPDVRPQPPREBEPPLPAARPAAGATLRRPKTLSPGKGVVQDVAPAGAVENPEYLPQ 1200
Qy 1201 GGAAPQPPRPAFPAFNDLYYMDPPERGAPESTFAGTPTAENPEYLGIDVPY 1255
Db 1201 GGAAPQPPRPAFPAFNDLYYMDPPERGAPESTFAGTPTAENPEYLGIDVPY 1255

```

RESULT 5

```

US-08-486-348A-68
; Sequence 68, Application US/08486348A
; Patent No. 5846538
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; APPLICATION NUMBER: US/08/486,348A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629

```

```

; REFERENCE/DOCKET NUMBER: 920010.448C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-486-348A-68
;
Query Match: 99.9%; Score 6806; DB 2; Length 1255;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
Qy 1 METALALCWGILLALLPFGAASVYCTGTDMKLRIPASEFTHLDMIRLUYQCGVQGNL 60
Db 1 METALALCWGILLALLPFGAASVYCTGTDMKLRIPASEFTHLDMIRLUYQCGVQGNL 60
Qy 61 ELTYLPTNASLFLQDIOEVQGVLIANNOYRQVPLQRLRIYRGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLFLQDIOEVQGVLIANNOYRQVPLQRLRIYRGTQLFEDNYALAVLDNG 120
Qy 121 DPLNNTTAVTGAAPGELRELQRLSLTEILKGVLIQRNPOLCYODTIILMKDIFHKNNQLA 180
Db 121 DPLNNTTAVTGAAPGELRELQRLSLTEILKGVLIQRNPOLCYODTIILMKDIFHKNNQLA 180
Qy 181 LTIIDTRNSRACHPSPCKSKSRCKGSESDQCILTRTYCAGGACRCKGPLETDCHEQC 240
Db 181 LTIIDTRNSRACHPSPCKSKSRCKGSESDQCILTRTYCAGGACRCKGPLETDCHEQC 240
Qy 241 AAGCTGPGHSDCLACIHNHSGICELHCPALVTYNTDFESMNPNEGRTFASCTYAC 300
Db 241 AAGCTGPGHSDCLACIHNHSGICELHCPALVTYNTDFESMNPNEGRTFASCTYAC 300
Qy 301 YNLTSTVGSCTLVCPHNOEVTABDGTORCEKSKPCARVYGIEMELREVRVATSN 360
Db 301 YNLTSTVGSCTLVCPHNOEVTABDGTORCEKSKPCARVYGIEMELREVRVATSN 360
Qy 361 IOEPAGCKKITIGSLAFLESTDGPASNTAPLOPQLOVFETLEITGYLYISAMPDSL 420
Db 361 IOEPAGCKKITIGSLAFLESTDGPASNTAPLOPQLOVFETLEITGYLYISAMPDSL 420
Qy 421 DLSVFQNLQVIRGILHNHAYSLTQGLISWLGIRSLREISGSLALHNHNLCPVHTV 480
Db 421 DLSVFQNLQVIRGILHNHAYSLTQGLISWLGIRSLREISGSLALHNHNLCPVHTV 480
Qy 481 PMDQLFNPQHALLHTANRPDEECVGEGLACHOLCARGHGWGPYTOCVNCSQFLRGEC 540
Db 481 PMDQLFNPQHALLHTANRPDEECVGEGLACHOLCARGHGWGPYTOCVNCSQFLRGEC 540
Qy 541 VESCRVIOGLPREYVNAHCLPCHPECOPONGSVTCFEPBADQCVACHYKDPFCVARC 600
Db 541 VESCRVIOGLPREYVNAHCLPCHPECOPONGSVTCFEPBADQCVACHYKDPFCVARC 600
Qy 601 PSQVPLSYMPKPEDEBGAQCPICNCTHSVDLDKCGCPAEQAPLTSISAVVG 660
Db 601 PSQVPLSYMPKPEDEBGAQCPICNCTHSVDLDKCGCPAEQAPLTSISAVVG 660
Qy 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOAMRIKETEL 720
Db 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOAMRIKETEL 720
Qy 721 RKVKVLGSGAFGTYYKGIWIPDGENVKIPVAIKYLRRENTSPKANKIIDEAYVAVGSP 780
Db 721 RKVKVLGSGAFGTYYKGIWIPDGENVKIPVAIKYLRRENTSPKANKIIDEAYVAVGSP 780
Qy 781 YVSRLLGICLTSTVOLVTOAMPYGCILDHVRENRGLSGODLNNWMOIAKMSYLEDR 840
Db 781 YVSRLLGICLTSTVOLVTOAMPYGCILDHVRENRGLSGODLNNWMOIAKMSYLEDR 840
Qy 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETRYHADGSKVPIKMMALSIILRRFT 900

```

Db 841 LVHRDLAARNVLKSPNNHYKITDGLARLLDIDETEHADGKVPKIMALLSILRRFT 900
 Qy 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIDPLEKEGRLPOPICTIDVYIMVKCM 960
 Db 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIDPLEKEGRLPOPICTIDVYIMVKCM 960
 Qy 961 IDSECRPRRELVSERFARMARDPQRFVVIQNEDLGPASPLDSTFYRSLLDDMDGLVDA 1020
 Db 961 IDSECRPRRELVSERFARMARDPQRFVVIQNEDLGPASPLDSTFYRSLLDDMDGLVDA 1020
 Qy 1021 BEYLVPQGFPCDPDPAAGAGWVHHRSSSTRSGGDLTLGLEPSEEBAPSLAPSEG 1080
 Db 1021 BEYLVPQGFPCDPDPAAGAGWVHHRSSSTRSGGDLTLGLEPSEEBAPSLAPSEG 1080
 Qy 1081 AGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPVPLPSETDGYVALTCSPOPEYV 1140
 Db 1081 AGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPVPLPSETDGYVALTCSPOPEYV 1140
 Qy 1141 NOPDVRPQPPSPREGPLPARPAGATLERPKTSPGKGVKQDVPAPGAVENPEYLTVQ 1200
 Db 1141 NOPDVRPQPPSPREGPLPARPAGATLERPKTSPGKGVKQDVPAPGAVENPEYLTVQ 1200
 Qy 1201 GGAAPQHPPPAPSPAPFNDLVYWDOPPRGAPSTFKGPTANPEYLGIDVPV 1255
 Db 1201 GGAAPQHPPPAPSPAPFNDLVYWDOPPRGAPSTFKGPTANPEYLGIDVPV 1255

RESULT 6
 US-08-468-545B-68
 / Sequence 68, Application US/08468545B
 / Patent No. 5876712

/ GENERAL INFORMATION:
 / APPLICANT: Cheever, Martin A.
 / APPLICANT: Delsis, Mary L.
 / TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
 / TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
 / TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
 / NUMBER OF SEQUENCES: 69
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Seed and Berry LLP
 / STREET: 6300 Columbia Center, 701 Fifth Avenue
 / CITY: Seattle
 / STATE: Washington
 / COUNTRY: US
 / ZIP: 98104-7092
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patentin Release #1.0, Version #1.25
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/468,545B
 / FILING DATE: 06-JUN-1995
 / CLASSIFICATION: 424
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Sharkey, Richard G.
 / REGISTRATION NUMBER: 32,629
 / REFERENCE/DOCKET NUMBER: 920010.448CS
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (206) 622-4900
 / TELEFAX: (206) 682-6031
 / INFORMATION FOR SEQ ID NO: 68:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1255 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / US-08-468-545B-68

Query Match 99.9%; Score 6806; DB 2; Length 1255;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 METALACRMGLLALLPAGASTOVCTGDMKRLPASPTHLDMLKRLHYOGCQVQGNL 60

Db 1 METALACRMGLLALLPAGASTOVCTGDMKRLPASPTHLDMLKRLHYOGCQVQGNL 60
 Qy 61 ELTYLPTNASSLFDIOIEVQGYVLIANQVQVPLQRLIRYRGTOPLFEDNVALAVLDNG 120
 Db 61 ELTYLPTNASSLFDIOIEVQGYVLIANQVQVPLQRLIRYRGTOPLFEDNVALAVLDNG 120
 Qy 121 DELNNTTPVTGASPGGLBELQRLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
 Db 121 DELNNTTPVTGASPGGLBELQRLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
 Qy 181 LTLIDTNSRACHPSPCKSGSRGSESDOSLTRTYCAGGACCKPLPTDCCHBOC 240
 Db 181 LTLIDTNSRACHPSPCKSGSRGSESDOSLTRTYCAGGACCKPLPTDCCHBOC 240
 Qy 241 AAGCTGPKASDCLACLFHNSGICELHCPALVTYNTDRESMPNPGRYTFGASCYTACP 300
 Db 241 AAGCTGPKASDCLACLFHNSGICELHCPALVTYNTDRESMPNPGRYTFGASCYTACP 300
 Qy 301 YNLTSTVGSCTLVCPRLNQEVTAEDGTQRCCKSPCARVCYGLGMEHLREYRAVTSAN 360
 Db 301 YNLTSTVGSCTLVCPRLNQEVTAEDGTQRCCKSPCARVCYGLGMEHLREYRAVTSAN 360
 Qy 361 IGEFAGCKKIFGSLAPLSPESFDGPASNTAPLOPEOLOVFEETLEITGYLISAMPDLP 420
 Db 361 IGEFAGCKKIFGSLAPLSPESFDGPASNTAPLOPEOLOVFEETLEITGYLISAMPDLP 420
 Qy 421 DLSEFQNLQVIRGRILHNGASVLTQGLGISWLGRLSRLBLSGLALIHNTHLCFVHTV 480
 Db 421 DLSEFQNLQVIRGRILHNGASVLTQGLGISWLGRLSRLBLSGLALIHNTHLCFVHTV 480
 Qy 481 PMDOLFNRPHQALLHTANRPDEBCVSGELACHQLCARHGCMGPGPOCVNCSQFLRQEC 540
 Db 481 PMDOLFNRPHQALLHTANRPDEBCVSGELACHQLCARHGCMGPGPOCVNCSQFLRQEC 540
 Qy 541 VECRVLQGLPREYVNAHCLCPHECPONGSVTCFGEADOCVCAHYKDPFCVARC 600
 Db 541 VECRVLQGLPREYVNAHCLCPHECPONGSVTCFGEADOCVCAHYKDPFCVARC 600
 Qy 601 PSGVKEPDLSPMPIMKPEDEGACQPCPINCTHSCVDLDKGCAPAEORASPLTIIISAVG 660
 Db 601 PSGVKEPDLSPMPIMKPEDEGACQPCPINCTHSCVDLDKGCAPAEORASPLTIIISAVG 660
 Qy 661 ILLVVLGVVFGILLIRROOKIRKTYMRLLQETELVEPLTPSGAMPNOAMILKETEL 720
 Db 661 ILLVVLGVVFGILLIRROOKIRKTYMRLLQETELVEPLTPSGAMPNOAMILKETEL 720
 Qy 721 RKVKVLGSGAFGVYKGIWIPGEENVKIPVALKVLBNTPSKANKETLDEAYVMAAGVSP 780
 Db 721 RKVKVLGSGAFGVYKGIWIPGEENVKIPVALKVLBNTPSKANKETLDEAYVMAAGVSP 780
 Qy 781 YNSRLIGICITSTVOLVTOAMPYGLLDHYENRGRGSGDILNMCQIAKMSYLEDDR 840
 Db 781 YNSRLIGICITSTVOLVTOAMPYGLLDHYENRGRGSGDILNMCQIAKMSYLEDDR 840
 Qy 841 LVHRDLAARNVLKSPNNHYKITDGLARLLDIDETEHADGKVPKIMALLSILRRFT 900
 Db 841 LVHRDLAARNVLKSPNNHYKITDGLARLLDIDETEHADGKVPKIMALLSILRRFT 900
 Qy 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIDPLEKEGRLPOPICTIDVYIMVKCM 960
 Db 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIDPLEKEGRLPOPICTIDVYIMVKCM 960
 Qy 961 IDSECRPRRELVSERFARMARDPQRFVVIQNEDLGPASPLDSTFYRSLLDDMDGLVDA 1020
 Db 961 IDSECRPRRELVSERFARMARDPQRFVVIQNEDLGPASPLDSTFYRSLLDDMDGLVDA 1020
 Qy 1021 BEYLVPQGFPCDPDPAAGAGWVHHRSSSTRSGGDLTLGLEPSEEBAPSLAPSEG 1080
 Db 1021 BEYLVPQGFPCDPDPAAGAGWVHHRSSSTRSGGDLTLGLEPSEEBAPSLAPSEG 1080
 Qy 1081 AGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPVPLPSETDGYVALTCSPOPEYV 1140

Db 1081 AGSDVDFDGLGMAKGLQSLPTHDPSPLOKRYSEDPVPLPSETDGVAVLTCSPQPEYV 1140
Qy 1141 NOPDVRRPQPSRREGGLPAARPAQATLTERPKTSLSPKNGVVKDVPAGGAVENPEYLTPO 1200
Db 1141 NOPDVRRPQPSRREGGLPAARPAQATLTERPKTSLSPKNGVVKDVPAGGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPSPADNLYYMDODPPERGAPSPSTFKGTFAENPEYLTGLDVPV 1255
Db 1201 GGAAPQHPHPAPSPADNLYYMDODPPERGAPSPSTFKGTFAENPEYLTGLDVPV 1255

RESULT 7

US-08-466-6808-68
; Sequence 68, Application US/084666808
; Patent No. 6075122
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Diehl, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,6808
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-466-6808-68

Query Match 99.9%; Score 6806; DB 3; Length 1255;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELAALCRWGLLALLPFGAASVQVCTGDMKLRLPASPETHLDMLRHLVYOGCCOVVQNL 60
Db 1 MELAALCRWGLLALLPFGAASVQVCTGDMKLRLPASPETHLDMLRHLVYOGCCOVVQNL 60
Qy 61 ELTYLPTNAGISLPIODIOEVQGYVLIANOVQVPLQSLRIVRGTQLPEDNYVALAVLDNG 120
Db 61 ELTYLPTNAGISLPIODIOEVQGYVLIANOVQVPLQSLRIVRGTQLPEDNYVALAVLDNG 120
Qy 121 DPLANNTPVTGASPGGLRELOLSITLTKGVLIOQRPOLCYDPTILMKDIFHKNNOLA 180
Db 121 DPLANNTPVTGASPGGLRELOLSITLTKGVLIOQRPOLCYDPTILMKDIFHKNNOLA 180
Qy 181 LTLIDTRSRACHPCSPWCKSRCKGSESDCQSLTRVCAAGGACRGLPTDCCHQOC 240
Db 181 LTLIDTRSRACHPCSPWCKSRCKGSESDCQSLTRVCAAGGACRGLPTDCCHQOC 240
Qy 241 AAGCTGRKSHSDCLACHFNHSGICELHCPALVTYNTDTFSMPNPEGRYTFGASCVTACP 300

Db 241 AAGCTGRKSHSDCLACHFNHSGICELHCPALVTYNTDTFSMPNPEGRYTFGASCVTACP 300
Qy 301 YNLTSTDVSCCTLVCCPLHNOETABEDGTQRCCKSKRCANVCIGLMEHLREVAATVSAN 360
Db 301 YNLTSTDVSCCTLVCCPLHNOETABEDGTQRCCKSKRCANVCIGLMEHLREVAATVSAN 360
Qy 361 IQEFACCKKIFGSLAFLPESFDGDPASNTAPLOPELOVETLEIETGYLYISAWPSLP 420
Db 361 IQEFACCKKIFGSLAFLPESFDGDPASNTAPLOPELOVETLEIETGYLYISAWPSLP 420
Qy 421 DISVFQNLQVIRGRILHNGAYSLTLQGLGISWLGSLRLRELSGLALIHNTHLCPFHVT 480
Db 421 DISVFQNLQVIRGRILHNGAYSLTLQGLGISWLGSLRLRELSGLALIHNTHLCPFHVT 480
Qy 481 PMDOLFRNPHOALLHNPANPEDECVESGLAQHLCARGHCWGSPPTQCVNCSOFLRQEC 540
Db 481 PMDOLFRNPHOALLHNPANPEDECVESGLAQHLCARGHCWGSPPTQCVNCSOFLRQEC 540
Qy 541 VEECRVLQGLPREVYNARCLPCHECOPONGSVTCFGEADQCVACAHYKDPFCVARC 600
Db 541 VEECRVLQGLPREVYNARCLPCHECOPONGSVTCFGEADQCVACAHYKDPFCVARC 600
Qy 601 PSQVPRDLSYMPIMKFPDEGACQPCPINCSTHSCVDLDDKGCAPADQVACAHYKDPFCVARC 660
Db 601 PSQVPRDLSYMPIMKFPDEGACQPCPINCSTHSCVDLDDKGCAPADQVACAHYKDPFCVARC 660
Qy 661 ILVVLGVVFGILLKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQOMRLKETEL 720
Db 661 ILVVLGVVFGILLKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQOMRLKETEL 720
Qy 721 RKVKYLGSGAGSTYKGIWIPGENVKIPVALKULRENTSPRANKIILDEAYMAGVSP 780
Db 721 RKVKYLGSGAGSTYKGIWIPGENVKIPVALKULRENTSPRANKIILDEAYMAGVSP 780
Qy 781 YVSRLLGICLTSTVOLVTOAMPYGCILDVRENRRGLSGQDILNMCMQIAKGSYLEYDR 840
Db 781 YVSRLLGICLTSTVOLVTOAMPYGCILDVRENRRGLSGQDILNMCMQIAKGSYLEYDR 840
Qy 841 LVHRDLAARNVLYKSPNNVKTIDFGALRLIDETERYHADGKVPKIMNALESILRRFT 900
Db 841 LVHRDLAARNVLYKSPNNVKTIDFGALRLIDETERYHADGKVPKIMNALESILRRFT 900
Qy 901 HOSDVMSYGVVWELMTGAKPYDGIIPAREIPDLLEKGERLPQPPICITDVMIMVCKWM 960
Db 901 HOSDVMSYGVVWELMTGAKPYDGIIPAREIPDLLEKGERLPQPPICITDVMIMVCKWM 960
Qy 961 IDSECRPRFRELVSFMSHARDPQRFVYIOWEDLGPASPILDSTFYRSILEDMDGDLVDA 1020
Db 961 IDSECRPRFRELVSFMSHARDPQRFVYIOWEDLGPASPILDSTFYRSILEDMDGDLVDA 1020
Qy 1021 EETLVPOQGFPCPPAPAGAGGVHRRHSSSTRSGGDLTLGLBPSSEBAPRSLAPSEB 1080
Db 1021 EETLVPOQGFPCPPAPAGAGGVHRRHSSSTRSGGDLTLGLBPSSEBAPRSLAPSEB 1080
Qy 1081 AGSDVDFDGLGMAKGLQSLPTHDPSPLOKRYSEDPVPLPSETDGVAVLTCSPQPEYV 1140
Db 1081 AGSDVDFDGLGMAKGLQSLPTHDPSPLOKRYSEDPVPLPSETDGVAVLTCSPQPEYV 1140
Qy 1141 NOPDVRRPQPSRREGGLPAARPAQATLTERPKTSLSPKNGVVKDVPAGGAVENPEYLTPO 1200
Db 1141 NOPDVRRPQPSRREGGLPAARPAQATLTERPKTSLSPKNGVVKDVPAGGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPSPADNLYYMDODPPERGAPSPSTFKGTFAENPEYLTGLDVPV 1255
Db 1201 GGAAPQHPHPAPSPADNLYYMDODPPERGAPSPSTFKGTFAENPEYLTGLDVPV 1255

RESULT 8
US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:

APPLICANT: Nicolette, Charles
 TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
 FILE REFERENCE: 126881309200
 CURRENT APPLICATION NUMBER: US/09/527,487
 CURRENT FILING DATE: 2000-03-16
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 2
 LENGTH: 1255
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-527-487-2

Query Match 99.9%; Score 6806; DB 4; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MELAALCRWGLLALPPGAASVQCTGDMKRLPASPEHLDMLRHLYQGQVQGNL 60
 1 MELAALCRWGLLALPPGAASVQCTGDMKRLPASPEHLDMLRHLYQGQVQGNL 60
 61 ELTYLPTNLSLFLQDIQEVQVYLIAHQVQVPLQRLRIYVGTOLFEEDNALAVLDNG 120
 61 ELTYLPTNLSLFLQDIQEVQVYLIAHQVQVPLQRLRIYVGTOLFEEDNALAVLDNG 120
 121 DPLANTPTVQASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILMKDITHKNNOLA 180
 121 DPLANTPTVQASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILMKDITHKNNOLA 180
 121 DPLANTPTVQASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILMKDITHKNNOLA 180
 181 LTLIDNRSBACHPCSPMCKGSRCWGESSSEDCSLRTVAGGACACGKPLPTDCHEOC 240
 181 LTLIDNRSBACHPCSPMCKGSRCWGESSSEDCSLRTVAGGACACGKPLPTDCHEOC 240
 241 AAGCTGPKGSDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCVTACP 300
 241 AAGCTGPKGSDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCVTACP 300
 301 YNTLSDVSGCTVLCPLANOETABDGTQCEKCSKPCAVCYGLGMEHLREVRATYSAN 360
 301 YNTLSDVSGCTVLCPLANOETABDGTQCEKCSKPCAVCYGLGMEHLREVRATYSAN 360
 361 IOEFACCKKIFGSLAPFESFDDPASPNTAPLOPEOLQVETLEBITGYLISAMPDLP 420
 361 IOEFACCKKIFGSLAPFESFDDPASPNTAPLOPEOLQVETLEBITGYLISAMPDLP 420
 421 DLSVPQNLQVIRGRILHNGAYSLTLQGLISWLGSLRELSGLALIHNTILCFVHTV 480
 421 DLSVPQNLQVIRGRILHNGAYSLTLQGLISWLGSLRELSGLALIHNTILCFVHTV 480
 481 PMDQLFRNPHQALLHTANREDECEVGEGLACHOLCARGHGCGPPTQCVNCSQFLRGQEC 540
 481 PMDQLFRNPHQALLHTANREDECEVGEGLACHOLCARGHGCGPPTQCVNCSQFLRGQEC 540
 541 VECRVLQGLPRRYNVARHCLPCHPECOPONGSVTCGPBPAADQVCAHAKDPFCVARC 600
 541 VECRVLQGLPRRYNVARHCLPCHPECOPONGSVTCGPBPAADQVCAHAKDPFCVARC 600
 601 PSGVKPDLSPYIWKPFDEGACQPCPINCSTHSCVDLDDGCPAPASPLTIIISAVNG 660
 601 PSGVKPDLSPYIWKPFDEGACQPCPINCSTHSCVDLDDGCPAPASPLTIIISAVNG 660
 661 ILLVVLGVVFGIILKRRQKIRKTYRRLLQETELVEPLTPSGAMPNOQMRLKETEL 720
 661 ILLVVLGVVFGIILKRRQKIRKTYRRLLQETELVEPLTPSGAMPNOQMRLKETEL 720
 721 RKVKLGSGAGVYVYKIMIPDGENYKIPVALIKYLRNTSPKANKELIDRAYVWAGVSP 780
 721 RKVKLGSGAGVYVYKIMIPDGENYKIPVALIKYLRNTSPKANKELIDRAYVWAGVSP 780
 781 YVSRLLGICLTSFVQVLTQMLPYGCLDHYENRGRISODLNMCMQIAKMSYLEDVR 840
 781 YVSRLLGICLTSFVQVLTQMLPYGCLDHYENRGRISODLNMCMQIAKMSYLEDVR 840

841 LVHRDLAARNVLKSPNHYKITDPLGLRLDIDETBYHADGGKVPKMMALBSILRRFT 900
 841 LVHRDLAARNVLKSPNHYKITDPLGLRLDIDETBYHADGGKVPKMMALBSILRRFT 900
 901 HQSDVMSYGVWMEIMTFGAKPYDGI PAREIDLEKGRLPQPICTIDVYIMVYKCM 960
 901 HQSDVMSYGVWMEIMTFGAKPYDGI PAREIDLEKGRLPQPICTIDVYIMVYKCM 960
 961 IDSECRPRERELVSERSMARDPORFVYQNBDLGPASPLDSTFFYSLSLEDDMGDLVDA 1020
 961 IDSECRPRERELVSERSMARDPORFVYQNBDLGPASPLDSTFFYSLSLEDDMGDLVDA 1020
 1021 EHYLVPOQGFPCDDPAPGAGVHHRHRSSTRSGGDLTLGLEPSEBAPRPLAPSEG 1080
 1021 EHYLVPOQGFPCDDPAPGAGVHHRHRSSTRSGGDLTLGLEPSEBAPRPLAPSEG 1080
 1081 AGSDVFDGLGMAKAGLQSLPTHPSPLORYSEDPVPLPSETDGVABLTCSPOREYV 1140
 1081 AGSDVFDGLGMAKAGLQSLPTHPSPLORYSEDPVPLPSETDGVABLTCSPOREYV 1140
 1141 NQPDVAPQPPSPREGLPAPAPAGATLERPKTISPKNGVYKVPAPGAVENPEYLTPO 1200
 1141 NQPDVAPQPPSPREGLPAPAPAGATLERPKTISPKNGVYKVPAPGAVENPEYLTPO 1200
 1201 GGAAPQHPPEPAPSPAFNDLYYWDOPPRBGAAPSTFGKPTAENPEYLGADVPV 1255
 1201 GGAAPQHPPEPAPSPAFNDLYYWDOPPRBGAAPSTFGKPTAENPEYLGADVPV 1255

RESULT 9
 US-08-484-438-8
 Sequence 8, Application US/08484438
 Patent No. 5811098
 Patent No. 5811098 5780031
 GENERAL INFORMATION:
 APPLICANT: Plowman, Gregory D.
 APPLICANT: Culoscon, Jean-Michel
 APPLICANT: Shoyab, Mohammed
 APPLICANT: Stegall, Clay B.
 APPLICANT: Heller m, Ingegerd
 APPLICANT: Heller m, Karl E.
 TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,438
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/323,442
 FILING DATE: 14-OCT-1994
 APPLICATION NUMBER: US 08/150,704
 FILING DATE: 10-NOV-1993
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/981,165
 FILING DATE: 24-NOV-1992
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Mierock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 5624-230

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIZ
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1255 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-484-438-8

Query Match 98.8%; Score 6732; DB 2; Length 1255;
 Best Local Similarity 99.3%; Pident. No. 0;
 Matches 1247; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

1 MELAALCRNGLLALALPPGAASVQVCTGDMKRLRLPASPETHLDMRLHYOGGVQGNL 60
 1 MELAALCRNGLLALALPPGAASVQVCTGDMKRLRLPASPETHLDMRLHYOGGVQGNL 60
 61 ELTYLPTNALSFLDIOEVQGVLLAHNOVQVPLQRLIYRGTLFEDNYALAVLDNG 120
 61 ELTYLPTNALSFLDIOEVQGVLLAHNOVQVPLQRLIYRGTLFEDNYALAVLDNG 120
 121 DPLANTPTTGASPGGLRELQRLSLEIIKGVLLIQNPOLCYQDTILWKDIFHKNOQLA 180
 121 DPLANTPTTGASPGGLRELQRLSLEIIKGVLLIQNPOLCYQDTILWKDIFHKNOQLA 180
 181 LTLIDNRSRACHPCSPCKSGRCSGSESDQSLRTVTCAGGCACCKPLPDDCCEOC 240
 181 LTLIDNRSRACHPCSPCKSGRCSGSESDQSLRTVTCAGGCACCKPLPDDCCEOC 240
 241 AAGCTGPKHSIDCLAFHNSGICELHCPALVTYNTDTPESNPPEGRYTFGASCVTAC 300
 241 AAGCTGPKHSIDCLAFHNSGICELHCPALVTYNTDTPESNPPEGRYTFGASCVTAC 300
 301 YNYLSTDVGSCTVCPRLHNOEYTAEDGTQRCCKSPKPCAVCYGLGMEHLREVAVTSAN 360
 301 YNYLSTDVGSCTVCPRLHNOEYTAEDGTQRCCKSPKPCAVCYGLGMEHLREVAVTSAN 360
 361 IOBFAGCKKIFGSLAFLEPSFGDDPASNTAPLOQVFEKTEBETGLYISAMPDILP 420
 361 IOBFAGCKKIFGSLAFLEPSFGDDPASNTAPLOQVFEKTEBETGLYISAMPDILP 420
 421 DLSVFNQLQVIRGRILHNGAYSLTLQGLISWLGRLSRLRELSGLALIHNNHLLCPVHTV 480
 421 DLSVFNQLQVIRGRILHNGAYSLTLQGLISWLGRLSRLRELSGLALIHNNHLLCPVHTV 480
 481 PWDOLFNNPQALLHTANRPEDCEVGEGLACHOLCARGHCMGPGPTOCVNCQSFARGQEC 540
 481 PWDOLFNNPQALLHTANRPEDCEVGEGLACHOLCARGHCMGPGPTOCVNCQSFARGQEC 540
 541 VEBGRVIOGLREVYNARHCLPCHPECCQPNQSVTCFGEADQCYACAHYKDPFCVARC 600
 541 VEBGRVIOGLREVYNARHCLPCHPECCQPNQSVTCFGEADQCYACAHYKDPFCVARC 600
 601 PSQVPPDLSYMPIMKFPDEEGACOPCPINCHTSCVDLDDKGPAPORASPLTISAVVG 660
 601 PSQVPPDLSYMPIMKFPDEEGACOPCPINCHTSCVDLDDKGPAPORASPLTISAVVG 660
 661 ILLVVLGVVFGIILKRROKIRKTYMRLLQETELVEPLTPSGAMPQAOQRILKETEL 720
 661 ILLVVLGVVFGIILKRROKIRKTYMRLLQETELVEPLTPSGAMPQAOQRILKETEL 720
 721 RKVVVLGGAGAGTYKGIWIPDGENVKIPVAIKVLRNTSPKANKIIDEAYVAVGSP 780
 721 RKVVVLGGAGAGTYKGIWIPDGENVKIPVAIKVLRNTSPKANKIIDEAYVAVGSP 780
 781 YVSHLLGICLSTVQLVTLMPYGLDHRVNRNGRLSGODLLMCMQIAGMSTLEVR 840
 781 YVSHLLGICLSTVQLVTLMPYGLDHRVNRNGRLSGODLLMCMQIAGMSTLEVR 840

QY 841 LVHEDLAARNVLKSPNHWKITDGLARLLDIDETBYHADGKVPKIMMALESILRRFT 900
 DB 841 LVHEDLAARNVLKSPNHWKITDGLARLLDIDETBYHADGKVPKIMMALESILRRFT 900
 QY 901 HQSDWVSYGVTWELMTFGAKPDGIPARIPDLLEGEERLPPOPICTIDVYMTMVKCM 960
 DB 901 HQSDWVSYGVTWELMTFGAKPDGIPARIPDLLEGEERLPPOPICTIDVYMTMVKCM 960
 QY 961 IDSECRPRRELVSERSMARDFORFVIONEDLGPASPLDSTFYRSLLEDMDGDLVDA 1020
 DB 961 IDSECRPRRELVSERSMARDFORFVIONEDLGPASPLDSTFYRSLLEDMDGDLVDA 1020
 QY 1021 ERTLVQOGFPCDPAPAGGVVHNRSSSTRSGGDLTLGLEPSEEA PRSPLAPSEG 1080
 DB 1021 ERTLVQOGFPCDPAPAGGVVHNRSSSTRSGGDLTLGLEPSEEA PRSPLAPSEG 1080
 QY 1081 AGSDVDFDGLMGCAAGLOSLPTHDPSPLOYSDEPTVPLPSSETDGYVAPLTCSPDEYV 1140
 DB 1081 AGSDVDFDGLMGCAAGLOSLPTHDPSPLOYSDEPTVPLPSSETDGYVAPLTCSPDEYV 1140
 QY 1141 NOPDVAPOPSPREGPLPARPAGATLERPKTLSPKNGVVKVAFAGAVENPEYLTPQ 1200
 DB 1141 NOPDVAPOPSPREGPLPARPAGATLERPKTLSPKNGVVKVAFAGAVENPEYLTPQ 1200
 QY 1201 GGAAPQHPPPAFSPAFDNLYYWDODPPRGAAPSTFKGPT-AENPEYGLDVPV 1255
 DB 1201 GGAAPQHPPPAFSPAFDNLYYWDODPPRGAAPSTFKGPT-AENPEYGLDVPV 1255

RESULT 10
 US-09-146-283-4
 ; Sequence 4, Application US/09146283
 ; Patent No. 5976546
 ; GENERAL INFORMATION:
 ; APPLICANT: Laus, Reiner
 ; APPLICANT: Ruegg, Curtis L.
 ; APPLICANT: Wu, Hongyu
 ; TITLE OF INVENTION: Immunostimulatory Compositions
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Denlinger & Associates
 ; STREET: 350 Cambridge Ave. Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/146.283
 ; FILING DATE: 03-SEPT-1998
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Judge, Linda R.
 ; REGISTRATION NUMBER: 42,702
 ; REFERENCE/DOCKET NUMBER: 7636-0010.21
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-324-0880
 ; TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 782 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: homo sapiens
 ; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
 ; US-09-146-283-4

Query Match 53.3%; Score 3630; DB 2; Length 782;
 Best Local Similarity 99.8%; Pred. No. 1.2e-238;
 Matches 653; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELALCRWGLLALPPGAASVQVCTGDMKRLPASPEHMLHLYOGCQVQGNL 60
DB 1 MELALCRWGLLALPPGAASVQVCTGDMKRLPASPEHMLHLYOGCQVQGNL 60
QY 61 ELYTLPTNASLSTLQDIQEVGYVLIANQVRYPLRLAIIVGTQLFEDNYALAVDNG 120
DB 61 ELYTLPTNASLSTLQDIQEVGYVLIANQVRYPLRLAIIVGTQLFEDNYALAVDNG 120
QY 121 DPLNNTTPTVGSFGSLRELQSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGSFGSLRELQSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLDITNRSBACHPCSPCKGSRGWSSESDQSLTRTVACAGGACRCKPLPTDCHEQC 240
DB 181 LTLDITNRSBACHPCSPCKGSRGWSSESDQSLTRTVACAGGACRCKPLPTDCHEQC 240
QY 241 AACCTGPKHSDDCLACHFNHSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYACP 300
DB 241 AACCTGPKHSDDCLACHFNHSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYACP 300
QY 301 YNLTSTDVGSCTIVCPILHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVAVTSAN 360
DB 301 YNLTSTDVGSCTIVCPILHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVAVTSAN 360
QY 361 IQEPAGCKKI FGSIAFLPESFDGDPASNTAPLOPEQLOVEETLEITGYLISAMPDLSL 420
DB 361 IQEPAGCKKI FGSIAFLPESFDGDPASNTAPLOPEQLOVEETLEITGYLISAMPDLSL 420
QY 421 DLSVFQNLQVIRGRILHNGAVSLTLOGISWGLSLRLSRELGSGLAIHNTHLCPVHTV 480
DB 421 DLSVFQNLQVIRGRILHNGAVSLTLOGISWGLSLRLSRELGSGLAIHNTHLCPVHTV 480
QY 481 PMDLFRNPHQALHTANRPEDCEVSGGLACHOLCARGHCWGPPTCCVNCISOPLRQEC 540
DB 481 PMDLFRNPHQALHTANRPEDCEVSGGLACHOLCARGHCWGPPTCCVNCISOPLRQEC 540
QY 541 VESCRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600
DB 541 VESCRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600
QY 601 PSQVKPDLSTMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCPRAGASPLTSL 654
DB 601 PSQVKPDLSTMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCPRAGASPLTSL 654

```

RESULT 11

US-08-579-823A-4
 Sequence 4, Application US/08579823A
 Patent No. 6080409

GENERAL INFORMATION:

APPLICANT: Laue, Reiner
 APPLICANT: Ruegg, Curdie L.
 APPLICANT: Wu, Hongyu
 TITLE OF INVENTION: Immunostimulatory Composition and Method
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Delinger & Associates
 STREET: 350 Cambridge Ave. Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/579, 823A
 FILING DATE: 03-DEC-1998
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Judge, Linda R.
 REGISTRATION NUMBER: 42,702
 REFERENCE/DOCKET NUMBER: 7636-0010
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 782 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: homo sapiens
 INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
 US-08-579-823A-4

Query Match 53.3%; Score 3630; DB 3; Length 782;
 Best Local Similarity 99.8%; Pred. No. 1.2e-238;
 Matches 653; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELALCRWGLLALPPGAASVQVCTGDMKRLPASPEHMLHLYOGCQVQGNL 60
DB 1 MELALCRWGLLALPPGAASVQVCTGDMKRLPASPEHMLHLYOGCQVQGNL 60
QY 61 ELYTLPTNASLSTLQDIQEVGYVLIANQVRYPLRLAIIVGTQLFEDNYALAVDNG 120
DB 61 ELYTLPTNASLSTLQDIQEVGYVLIANQVRYPLRLAIIVGTQLFEDNYALAVDNG 120
QY 121 DPLNNTTPTVGSFGSLRELQSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGSFGSLRELQSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLDITNRSBACHPCSPCKGSRGWSSESDQSLTRTVACAGGACRCKPLPTDCHEQC 240
DB 181 LTLDITNRSBACHPCSPCKGSRGWSSESDQSLTRTVACAGGACRCKPLPTDCHEQC 240
QY 241 AACCTGPKHSDDCLACHFNHSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYACP 300
DB 241 AACCTGPKHSDDCLACHFNHSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYACP 300
QY 301 YNLTSTDVGSCTIVCPILHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVAVTSAN 360
DB 301 YNLTSTDVGSCTIVCPILHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVAVTSAN 360
QY 361 IQEPAGCKKI FGSIAFLPESFDGDPASNTAPLOPEQLOVEETLEITGYLISAMPDLSL 420
DB 361 IQEPAGCKKI FGSIAFLPESFDGDPASNTAPLOPEQLOVEETLEITGYLISAMPDLSL 420
QY 421 DLSVFQNLQVIRGRILHNGAVSLTLOGISWGLSLRLSRELGSGLAIHNTHLCPVHTV 480
DB 421 DLSVFQNLQVIRGRILHNGAVSLTLOGISWGLSLRLSRELGSGLAIHNTHLCPVHTV 480
QY 481 PMDLFRNPHQALHTANRPEDCEVSGGLACHOLCARGHCWGPPTCCVNCISOPLRQEC 540
DB 481 PMDLFRNPHQALHTANRPEDCEVSGGLACHOLCARGHCWGPPTCCVNCISOPLRQEC 540
QY 541 VESCRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600
DB 541 VESCRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600
QY 601 PSQVKPDLSTMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCPRAGASPLTSL 654
DB 601 PSQVKPDLSTMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCPRAGASPLTSL 654

```

RESULT 12

US-09-344-195-4

Sequence 4, Application US/09344195
Patent No. 6210662
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
Ruegg, Curtis L.
Mu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,195
FILING DATE: 24-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0960
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-344-195-4
Query Match 53.3%; Score 3630; DB 3; Length 782;
Best Local Similarity 99.8%; Pred. No. 1,2e-238;
Matches 653; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLALPPGASTQVCTGTDMLRLPASPEHLDMLRHLYOGQGVVQNL 60
DB 1 MELAALCRWGLLALPPGASTQVCTGTDMLRLPASPEHLDMLRHLYOGQGVVQNL 60
QY 61 ELTYIPTNASLSPLODIOEVQGVILAHNOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
DB 61 ELTYIPTNASLSPLODIOEVQGVILAHNOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
QY 121 DELNNTTPTVTSRPGSLRELQRLSLREILKGVILIRNQLCYODITLWKDIFHKNNQLA 180
DB 121 DELNNTTPTVTSRPGSLRELQRLSLREILKGVILIRNQLCYODITLWKDIFHKNNQLA 180
QY 181 LTLIDTNSRACHPCSPCKSGRCWGESBDCSLTRTVCAAGCARCKGFLPTDCCHQC 240
DB 181 LTLIDTNSRACHPCSPCKSGRCWGESBDCSLTRTVCAAGCARCKGFLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFESMPNDEGRVTPGASCTYAC 300
DB 241 AAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFESMPNDEGRVTPGASCTYAC 300
QY 301 YNYLSTDVGSCTLVCPHNOETVTAEDGTQRCCKSKPCARVCYGLGNEHLREVRVTSAN 360
DB 301 YNYLSTDVGSCTLVCPHNOETVTAEDGTQRCCKSKPCARVCYGLGNEHLREVRVTSAN 360

DB 301 YNYLSTDVGSCTLVCPHNOETVTAEDGTQRCCKSKPCARVCYGLGNEHLREVRVTSAN 360
QY 361 IOEFAGCKKIFGSLAFLPESFDDPASNTAPLOPEQLOVETLEITGYLYISAMPDLP 420
DB 361 IOEFAGCKKIFGSLAFLPESFDDPASNTAPLOPEQLOVETLEITGYLYISAMPDLP 420
QY 421 DISVFONLQVTRRIILHNGVSLTLOGLISWLGRLSRLGSLALIHNTHLCPVHTV 480
DB 421 DISVFONLQVTRRIILHNGVSLTLOGLISWLGRLSRLGSLALIHNTHLCPVHTV 480
QY 481 PMQLFRNPHQALHTANRPEDSCVGBGLAQHLCARGHGWGSPPTQCVNCSQFLRGQEC 540
DB 481 PMQLFRNPHQALHTANRPEDSCVGBGLAQHLCARGHGWGSPPTQCVNCSQFLRGQEC 540
QY 541 VEECRVLOGLPREYVNAHRLCPCHPECOPNGSVTCGFPADQCVACAHYKDPFCVARC 600
DB 541 VEECRVLOGLPREYVNAHRLCPCHPECOPNGSVTCGFPADQCVACAHYKDPFCVARC 600
QY 601 PSCVAPDLSYMPIWKFPEDEGACQPCPINCTHSCVDLDDKGCFAEORASPLTSI 654
DB 601 PSCVAPDLSYMPIWKFPEDEGACQPCPINCTHSCVDLDDKGCFAEORASPLTSI 654
RESULT 13
US-08-422-108-1
Sequence 1, Application US/08422108
Patent No. 6015567
GENERAL INFORMATION:
APPLICANT: Hudzik, Robert M.
APPLICANT: Shepard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,108
FILING DATE: 14-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/355460
FILING DATE: 13-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/048346
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/354319
FILING DATE: 19-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554C2D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
US-08-422-108-1

Query Match 51.0%; Score 3473; DB 3; Length 624;
 Best Local Similarity 99.8%; Pred. No. 4.2e-228;
 Matches 623; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 22 STQVCTGTMKRLPASPETHLMDLRLHYGCGVVOGNIETLYLPTNASISFLQDIOEVQ 81
DB 1 STQVCTGTMKRLPASPETHLMDLRLHYGCGVVOGNIETLYLPTNASISFLQDIOEVQ 60
QY 82 GYVLIHNVQVQPLQRLRIVRGTOLEFEDNYALAVLDNGPLNNTPVTGASPGGLRELO 141
DB 61 GYVLIHNVQVQPLQRLRIVRGTOLEFEDNYALAVLDNGPLNNTPVTGASPGGLRELO 120
QY 142 LNSLTELKGGVLIQGNPOLCYODTILMKDIFKNNQALATLIDTNSRACHPCSPCKG 201
DB 121 LNSLTELKGGVLIQGNPOLCYODTILMKDIFKNNQALATLIDTNSRACHPCSPCKG 180
QY 202 SRMGSSSEDCQSLRTVCAAGGACRCKGPLPTDCHEQCAAGCTGPRHSDCLCLHFNHS 261
DB 181 SRMGSSSEDCQSLRTVCAAGGACRCKGPLPTDCHEQCAAGCTGPRHSDCLCLHFNHS 240
QY 262 GICELHCPALVTYNTDTFESMPNBEGRYTFGASCVTACPNYISTDVGCTLVCPPLHNOE 321
DB 241 GICELHCPALVTYNTDTFESMPNBEGRYTFGASCVTACPNYISTDVGCTLVCPPLHNOE 300
QY 322 VTAEDETQRCCKSKPCARVCYGLGMEHLREVAVTSANTIOEPAGCKKIFGSLAFLEPSF 381
DB 301 VTAEDETQRCCKSKPCARVCYGLGMEHLREVAVTSANTIOEPAGCKKIFGSLAFLEPSF 360
QY 382 DQDPASNTAPLOEPOLOVFETLEITGYLIASMPDLPPLSVFQNLQVIRGRILHNGAY 441
DB 361 DQDPASNTAPLOEPOLOVFETLEITGYLIASMPDLPPLSVFQNLQVIRGRILHNGAY 420
QY 442 SLTLQGLISWLGRLSRLREISGLALIHNTHLCFVHTVPMDOLEFRPHQALLHTANRPE 501
DB 421 SLTLQGLISWLGRLSRLREISGLALIHNTHLCFVHTVPMDOLEFRPHQALLHTANRPE 480
QY 502 DRCVGBGLACHOLCARGHCWGPPTQCVNCSQPLRGQECYBECHVLOGLEPREYVNAHCL 561
DB 481 DRCVGBGLACHOLCARGHCWGPPTQCVNCSQPLRGQECYBECHVLOGLEPREYVNAHCL 540
QY 562 PCHPBECOPNGSVTCFGEPAADQCAAHYKDPFCVACRCSGVKPDLSYMPIMKFPDEEG 621
DB 541 PCHPBECOPNGSVTCFGEPAADQCAAHYKDPFCVACRCSGVKPDLSYMPIMKFPDEEG 600
QY 622 ACQPCPINCTHSCVLDLDDKCPAE 645
DB 601 ACQPCPINCTHSCVLDLDDKCPAE 624

```

RESULT 14
 US-08-422-734-1
 Sequence 1, Application US/08422734
 Patent No. 6333169
 GENERAL INFORMATION:
 APPLICANT: Huddiak, Robert M.
 APPLICANT: Shepard, H. Michael
 APPLICANT: Ullrich, Axel
 TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatIn (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/422,734

FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/422108
 FILING DATE: 14-APR-1995
 APPLICATION NUMBER: 08/355460
 FILING DATE: 13-DEC-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/048346
 FILING DATE: 15-APR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/354319
 FILING DATE: 19-MAY-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Lee, Wendy M
 REGISTRATION NUMBER: 00,000
 REFERENCE/DOCKET NUMBER: 554C201
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-1994
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 624 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear

Query Match 51.0%; Score 3473; DB 4; Length 624;
 Best Local Similarity 99.8%; Pred. No. 4.2e-228;
 Matches 623; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 22 STQVCTGTMKRLPASPETHLMDLRLHYGCGVVOGNIETLYLPTNASISFLQDIOEVQ 81
DB 1 STQVCTGTMKRLPASPETHLMDLRLHYGCGVVOGNIETLYLPTNASISFLQDIOEVQ 60
QY 82 GYVLIHNVQVQPLQRLRIVRGTOLEFEDNYALAVLDNGPLNNTPVTGASPGGLRELO 141
DB 61 GYVLIHNVQVQPLQRLRIVRGTOLEFEDNYALAVLDNGPLNNTPVTGASPGGLRELO 120
QY 142 LNSLTELKGGVLIQGNPOLCYODTILMKDIFKNNQALATLIDTNSRACHPCSPCKG 201
DB 121 LNSLTELKGGVLIQGNPOLCYODTILMKDIFKNNQALATLIDTNSRACHPCSPCKG 180
QY 202 SRMGSSSEDCQSLRTVCAAGGACRCKGPLPTDCHEQCAAGCTGPRHSDCLCLHFNHS 261
DB 181 SRMGSSSEDCQSLRTVCAAGGACRCKGPLPTDCHEQCAAGCTGPRHSDCLCLHFNHS 240
QY 262 GICELHCPALVTYNTDTFESMPNBEGRYTFGASCVTACPNYISTDVGCTLVCPPLHNOE 321
DB 241 GICELHCPALVTYNTDTFESMPNBEGRYTFGASCVTACPNYISTDVGCTLVCPPLHNOE 300
QY 322 VTAEDETQRCCKSKPCARVCYGLGMEHLREVAVTSANTIOEPAGCKKIFGSLAFLEPSF 381
DB 301 VTAEDETQRCCKSKPCARVCYGLGMEHLREVAVTSANTIOEPAGCKKIFGSLAFLEPSF 360
QY 382 DQDPASNTAPLOEPOLOVFETLEITGYLIASMPDLPPLSVFQNLQVIRGRILHNGAY 441
DB 361 DQDPASNTAPLOEPOLOVFETLEITGYLIASMPDLPPLSVFQNLQVIRGRILHNGAY 420
QY 442 SLTLQGLISWLGRLSRLREISGLALIHNTHLCFVHTVPMDOLEFRPHQALLHTANRPE 501
DB 421 SLTLQGLISWLGRLSRLREISGLALIHNTHLCFVHTVPMDOLEFRPHQALLHTANRPE 480
QY 502 DRCVGBGLACHOLCARGHCWGPPTQCVNCSQPLRGQECYBECHVLOGLEPREYVNAHCL 561
DB 481 DRCVGBGLACHOLCARGHCWGPPTQCVNCSQPLRGQECYBECHVLOGLEPREYVNAHCL 540
QY 562 PCHPBECOPNGSVTCFGEPAADQCAAHYKDPFCVACRCSGVKPDLSYMPIMKFPDEEG 621
DB 541 PCHPBECOPNGSVTCFGEPAADQCAAHYKDPFCVACRCSGVKPDLSYMPIMKFPDEEG 600
QY 622 ACQPCPINCTHSCVLDLDDKCPAE 645

```

Db 601 ACOPPCINCHSCVDLDDKCCPAB 624

RESULT 15

US-08-475-035-4
Sequence 4, Application US/08475035
Patent No. 5985553

GENERAL INFORMATION:

APPLICANT: KING, C. R.
APPLICANT: KRAUS, MATTHIAS H.
APPLICANT: AARONSON, STUART A.
TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
TITLE OF INVENTION: EGF RECEPTOR GENE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: NEEDLE & ROSENBERG, P.C.
STREET: Suite 1200, 127 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475, 035
FILING DATE: 7 Jun 1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414.656
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 1210 amino acids
TYPE: amino acid
TOPOLOGY: 1linear
MOLECULE TYPE: protein

US-08-475-035-4

Query Match 46.5%; Score 3169; DB 2; Length 1210;
Best Local Similarity 49.8%; Pred. No. 4, 5e-207;
Matches 632; Conservative 177; Mismatches 345; Indels 116; Gaps 22;

QY 11 LLLALLPPGAA--STOVCTGTDMKRLTPASPTHLDMLEHLVYOGCOVGNLELYLPTN 68
DB 14 LLLALLPPGAA--STOVCTGTDMKRLTPASPTHLDMLEHLVYOGCOVGNLELYLPTN 73
QY 69 ASLSLTDIOYQGYVLAHNOVROPLQRLIVRGTOLEFDNALAVLNDGDPPLNNTTP 128
DB 74 YDLSEFLKIOYAGVYLAHNOVROPLQRLIVRGTOLEFDNALAVLNDGDPPLNNTTP 126
QY 129 VTGASPGSLRELOLSLLEILLKGVLIQNPOLCTODITLMDITLHKNQALTLIDNR 188
DB 127 ---ANKTGLKEIPMNLLOEILLHGAVERFSNNPALCNVESTIQMRDIVSSPFLSNMSDFQNH 183
QY 189 SRACHPCSPMGKSHCMWSESSDCSLTRTVCAAGCA--RCKGPLPTDCCHEGCAAGCTGP 247
DB 184 LGSQCKDPSCPNGSCWAGGECCKLTKIICAQCCSGCRCKSKSDCCCHNCAAGCTGP 243
QY 248 KHSQDLALHNRHSICELHCPALVTYNTDFESMNPBGRATYGASCCTACPYNYLSTD 307
DB 244 RESQDLVCRKFRDEATKCTCPPLMLYNPTTYQMDVNPEGKYSFGATCVKCKPRYVYVTD 303
QY 308 VGSCTIVPLINQETATADGTCORCEKSKPCKARVYCYGLGMEHLREVRATVTSANIQEPAGC 367
DB 304 HGSCTVACGADSYEM--BEDGVAKCKCKGCPCKVNCVIGIGIFKDSLSINATINIGIFKNC 362

QY 368 KTFISGLAFLPSFDDDPASNTAPLOPEOLQVEFTELEITGYLYISAMPDPLPSYFON 427
DB 363 TSIISDGLHLPLVAFRDSFTHTPPLDPELDIKYKEITGFLPLQAMPERTDLHAFEN 422
QY 428 LOYIRGRILHNGAVSLTLOGLSISGLSRLSGGLAIHNHTLCPVHTVWMDLFR 487
DB 423 LEIIRGRTOHQGFSLAVVSLNITSLSLRSLSKISDGDVILSGNKNLCYANTIMWKLFG 482
QY 488 NPHQALHTANPEDECSVEGLACHOLCARHGWGCPQCVNCSQSLRQCEBEERVL 547
DB 483 TSGQTKIISNGENSCATGVCYCHALCPBEGWGEPRDVCSCRVSRRRECVDKKLT 542
QY 548 OGLPREYVNAHCLPCHPECOPONGSVTCFGEPAQCAVCAHYKDPFCVAPRCSPGKPD 607
DB 543 EGPREFVENSECICHPRECLDQAMNITCTGRPNDCICCAHIIDGPHCVKCTPAGWGE 602
QY 608 LSYPLWKFPEDEGACQPCPINCCHSCVDLDDKCPAQRASPLTJISAVG--LLV 664
DB 603 NNML--VMKYADAGHVCMLCHPNCTYCTGPGLEGCPNGPKIP--SIATGMVGLLLLV 659
QY 665 VTLGVVFGILLIRROOKIRKTYMRLLQTELEVPLTBSGAMPNOQMRILKTELKVK 724
DB 660 VALGIG--LFMRRRHIVAKRLRLRLORELEVPLTBSGAMPNOQMRILKTELKVK 716
QY 725 VLSGAFGTVYKGIWIPDENYKIPVAIKVLRNTPSKANKETLDEAYVAVGSPVSR 784
DB 721 VLSGAFGTVYKGIWIPDENYKIPVAIKVLRNTPSKANKETLDEAYVAVGSPVSR 776
QY 785 LIGICLSTVQQLVTOAMPYGLLDHYRENRGLSGDLMQCAKAGSYLDEVLVHR 844
DB 777 ILGICLSTVQQLVTOAMPYGLLDHYRENRGLSGDLMQCAKAGSYLDEVLVHR 836
QY 845 DLAAENVLVKSPNHYKITDPLARLIDETVHADGCVPIKMALESILRRRFTHQSD 904
DB 837 DLAAENVLVKSPNHYKITDPLARLIDETVHADGCVPIKMALESILRRRFTHQSD 896
QY 905 VMSYGVYVWELMTFPAKPYDGIPIAREIPDLLEKGERLPOPICTIDVYIMVYCMNIDSE 964
DB 897 VMSYGVYVWELMTFPAKPYDGIPIAREIPDLLEKGERLPOPICTIDVYIMVYCMNIDSE 956
QY 965 CRPRELVSEFSRNARDPQRFVVIQ--NBDLQPASPLDSTFYRSLDEDDMDGLVDAEY 1023
DB 957 SRPKRELILIEFSKVARDPQRYLVIQSDBRMHLPSPTDSNFRALMDEEDMDVDAEY 1016
QY 1024 LVPQGFPCPDPAFGAGMHHRRHSSSTRSGGDLTLGLEPSEEARPSPLAPSGAGS 1083
DB 1017 LIPQGFPCPDPAFGAGMHHRRHSSSTRSGGDLTLGLEPSEEARPSPLAPSGAGS 1042
QY 1084 DVFDGDLGMAKGLQSLPTHDPSLQRYSDPTVPLPSRT--DGVAVPLTCSPOPEYVN 1141
DB 1043 N--NSTVACIDRNGAQSCPIYEDSFLQRYSDPTALTEDSIDDTFL-----PPEYIN 1094
QY 1142 QPDVAPPOPSPREGLPAPAPAGATL-----ERPFTLSPGKGVAVKDVAFGAGVAPENY 1196
DB 1095 Q-----SVF--KRAAGSVQNPVYHNPQPLNPAAPSRPHQD--PHSTAVGNPEY 1138
QY 1197 L--TPOGCAAPPHPPAPSPAPNDLYWDQ-----DP-----PERGAPSTFGK 1239
DB 1139 LNTVQ-----PTCVNSTFDSAPMAQKSHQISLDNPDYQDDFFPKAKPNCIGIFK 1189
QY 1240 TPTAENPEYL 1249
DB 1190 S-TAENPEYL 1198

Search completed: February 9, 2004, 16:17:57
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2004, 13:30:29 ; Search time 34 Seconds

(without alignments)
61.583 Million cell updates/sec

Title: US-09-930-125-3

Perfect score: 54

Sequence: 1 EBYLVPQGGF 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	10	US-09-930-125-3	Sequence 3, Appl1
2	54	100.0	10	US-10-313-644-3	Sequence 3, Appl1
3	54	100.0	59	US-09-854-356-5	Sequence 5, Appl1
4	54	100.0	266	US-09-854-356-7	Sequence 4, Appl1
5	54	100.0	293	US-10-102-806-583	Sequence 583, Appl1
6	54	100.0	570	US-10-378-393-15	Sequence 15, Appl1
7	54	100.0	583	US-09-930-125-9	Sequence 9, Appl1
8	54	100.0	589	US-09-930-125-8	Sequence 8, Appl1
9	54	100.0	589	US-09-930-125-10	Sequence 10, Appl1
10	54	100.0	600	US-09-930-125-11	Sequence 11, Appl1
11	54	100.0	712	US-09-854-356-7	Sequence 7, Appl1
12	54	100.0	791	US-10-378-393-11	Sequence 11, Appl1
13	54	100.0	919	US-09-854-356-6	Sequence 6, Appl1
14	54	100.0	1253	US-10-146-473-72	Sequence 72, Appl1
15	54	100.0	1255	US-09-811-123-9	Sequence 9, Appl1

16	54	100.0	1255	9	US-09-811-115-3	Sequence 3, Appl1
17	54	100.0	1255	10	US-09-769-508-2	Sequence 2, Appl1
18	54	100.0	1255	10	US-09-854-356-1	Sequence 1, Appl1
19	54	100.0	1255	10	US-09-930-125-2	Sequence 2, Appl1
20	54	100.0	1255	11	US-09-441-411-6	Sequence 6, Appl1
21	54	100.0	1255	12	US-10-207-498-6	Sequence 6, Appl1
22	54	100.0	1255	12	US-10-338-730-2	Sequence 2, Appl1
23	54	100.0	1255	12	US-10-313-644-2	Sequence 2, Appl1
24	54	100.0	1255	12	US-10-322-892-4	Sequence 2, Appl1
25	54	100.0	1255	12	US-10-272-437A-28	Sequence 28, Appl1
26	54	100.0	1255	12	US-10-117-937A-59A	Sequence 59A, Appl1
27	54	100.0	1255	12	US-10-394-322A-17	Sequence 17, Appl1
28	54	100.0	1255	12	US-10-245-871-553	Sequence 553, Appl1
29	54	100.0	1255	12	US-10-149-138-4641	Sequence 4641, Appl1
30	54	100.0	1255	12	US-10-435-696-36	Sequence 36, Appl1
31	54	100.0	1255	15	US-10-207-655-45	Sequence 45, Appl1
32	54	100.0	1255	15	US-10-177-293-126	Sequence 126, Appl1
33	54	100.0	1256	10	US-09-854-356-2	Sequence 2, Appl1
34	54	100.0	1256	10	US-09-854-356-14	Sequence 14, Appl1
35	54	100.0	1260	10	US-09-870-759-118	Sequence 118, Appl1
36	54	100.0	1260	12	US-09-751-708A-118	Sequence 118, Appl1
37	50	92.6	20	12	US-10-238-960-5	Sequence 5, Appl1
38	50	92.6	20	12	US-10-238-960-6	Sequence 6, Appl1
39	50	92.6	1210	9	US-09-725-433-2	Sequence 2, Appl1
40	50	92.6	1210	12	US-10-394-322A-16	Sequence 16, Appl1
41	49	90.7	9	12	US-10-245-871-607	Sequence 607, Appl1
42	49	90.7	9	12	US-10-149-138-1138	Sequence 1138, Appl1
43	49	90.7	9	12	US-10-149-138-3666	Sequence 3666, Appl1
44	49	90.7	10	12	US-10-149-138-1139	Sequence 1139, Appl1
45	49	90.7	10	12	US-10-149-138-3667	Sequence 3667, Appl1

ALIGNMENTS

RESULT 1
US-09-930-125-3
; Sequence 3, Application US/09930125
; Publication No. US20020193329A1
GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-125-3
Query Match 100.0%; Score 54; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00068;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 EBYLVPQGGF 10
Db 1 EBYLVPQGGF 10
RESULT 2
US-10-313-644-3
; Sequence 3, Application US/10313644
; Publication No. US20030157119A1

GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Cheever, Martin A.
APPLICANT: Hand-Zimmerman, Susan
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
FILE REFERENCE: 210121.483C3
CURRENT APPLICATION NUMBER: US/10/313,644
CURRENT FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 3
LENGTH: 10
TYPE: PRT
ORGANISM: Homo sapiens
US-10-313-644-3

Query Match 100.0%; Score 54; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10
DB 1 EBYLVPQGGF 10

RESULT 3
US-09-854-356-5
Sequence 5; Application US/09854356
Patent No. US2002017567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Cheever, Dirk
APPLICANT: Corixa Corporation
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356
CURRENT FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: US 09/493,480
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,976
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
OTHER INFORMATION: portion (delta PD) of human HER-2/neu
US-09-854-356-5

Query Match 100.0%; Score 54; DB 10; Length 59;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10
DB 32 EBYLVPQGGF 41

RESULT 4
US-09-854-356-4
Sequence 4; Application US/09854356
Patent No. US2002017567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Cheever, Dirk
APPLICANT: Corixa Corporation
TITLE OF INVENTION: HER-2/neu Fusion Proteins

FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356
CURRENT FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: US 09/493,480
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,976
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 266
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
US-09-854-356-4

Query Match 100.0%; Score 54; DB 10; Length 266;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10
DB 32 EBYLVPQGGF 41

RESULT 5
US-10-102-806-583
Sequence 583; Application US/10102806
Publication No. US2003005442A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103PIC1
CURRENT APPLICATION NUMBER: US/10/102,806
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/925,298
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 583
LENGTH: 293
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (52)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (53)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (58)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (150)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (156)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (171)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (207)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (254)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-583

Query Match 100.0%; Score 54; DB 15; Length 293;

Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 105 BEYLVPOQGF 114

RESULT 6
US-10-378-393-15

/ Sequence 15, Application US/10378393
/ Publication No. US20030182668A1
/ GENERAL INFORMATION:

/ APPLICANT: Bol, David K.
/ APPLICANT: Cardoni, Joan M.
/ APPLICANT: Rowley, Ronald B.
/ APPLICANT: Wong, Tai W.
/ APPLICANT: Lee, Francis Y.

/ TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
/ FILE REFERENCE: D0254 NP

/ CURRENT APPLICATION NUMBER: US/10/378,393
/ CURRENT FILING DATE: 2003-03-03

/ PRIOR APPLICATION NUMBER: US 60/360,889
/ PRIORITY FILING DATE: 2002-03-01

/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: Patent version 3.2

/ SEQ ID NO 15
/ LENGTH: 570

/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-378-393-15

Query Match 100.0%; Score 54; DB 12; Length 570;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 336 BEYLVPOQGF 345

RESULT 7
US-09-930-125-9

/ Sequence 9, Application US/09930125
/ Publication No. US20020193329A1
/ GENERAL INFORMATION:

/ APPLICANT: Hand-Zimmerman, Susan
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Foy, Teresa M.

/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: McNeill, Patricia D.

/ APPLICANT: Vedvick, Thomas S.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS

/ FILE REFERENCE: 210121.544
/ CURRENT APPLICATION NUMBER: US/09/930,125
/ CURRENT FILING DATE: 2001-08-14

/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0

/ SEQ ID NO 9
/ LENGTH: 583

/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-930-125-9

Query Match 100.0%; Score 54; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 105 BEYLVPOQGF 114

Db 347 BEYLVPOQGF 356

RESULT 8
US-09-930-125-8

/ Sequence 8, Application US/09930125
/ Publication No. US20020193329A1
/ GENERAL INFORMATION:

/ APPLICANT: Hand-Zimmerman, Susan
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Foy, Teresa M.

/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: McNeill, Patricia D.

/ APPLICANT: Vedvick, Thomas S.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS

/ FILE REFERENCE: 210121.544
/ CURRENT APPLICATION NUMBER: US/09/930,125
/ CURRENT FILING DATE: 2001-08-14

/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0

/ SEQ ID NO 8
/ LENGTH: 587

/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-930-125-8

Query Match 100.0%; Score 54; DB 10; Length 587;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 347 BEYLVPOQGF 356

RESULT 9
US-09-930-125-10

/ Sequence 10, Application US/09930125
/ Publication No. US20020193329A1
/ GENERAL INFORMATION:

/ APPLICANT: Hand-Zimmerman, Susan
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Foy, Teresa M.

/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: McNeill, Patricia D.

/ APPLICANT: Vedvick, Thomas S.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS

/ FILE REFERENCE: 210121.544
/ CURRENT APPLICATION NUMBER: US/09/930,125
/ CURRENT FILING DATE: 2001-08-14

/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0

/ SEQ ID NO 10
/ LENGTH: 589

/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-930-125-10

Query Match 100.0%; Score 54; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 355 BEYLVPOQGF 364

RESULT 10
US-09-930-125-11


```
/ Sequence 11, Application US/09930125
/ Publication No. US20020193329A1
/ GENERAL INFORMATION:
/ APPLICANT: Hand-Zimmerman, Susan
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Roy, Teresa M.
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Vedvick, Thomas S.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
/ TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
/ FILE REFERENCE: 210121.544
/ CURRENT APPLICATION NUMBER: US/09/930.125
/ CURRENT FILING DATE: 2001-08-14
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 600
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-930-125-11

Query Match      100.0%; Score 54; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
Db 366 EBYLVPOQGF 375

RESULT 11
US-09-854-356-7
/ Sequence 7, Application US/09854356
/ Patent No. US2002017567A1
/ GENERAL INFORMATION:
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Gheysen, Dirk
/ APPLICANT: Corixa Corporation
/ APPLICANT: SmithKline Beecham Biologicals S. A.
/ TITLE OF INVENTION: HER-2/neu Fusion Proteins
/ FILE REFERENCE: 014058-009810PC
/ CURRENT APPLICATION NUMBER: US/09/854.356
/ CURRENT FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 09/493,480
/ PRIOR FILING DATE: 2000-01-28
/ PRIOR APPLICATION NUMBER: US 60/117,976
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 7
/ LENGTH: 712
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:fusion protein
/ OTHER INFORMATION: of ECD and delta PD of human HER-2/neu
US-09-854-356-7

Query Match      100.0%; Score 54; DB 10; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
Db 685 EBYLVPOQGF 694

RESULT 12
US-10-378-393-11
/ Sequence 11, Application US/10378393
/ Publication No. US2003018268A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Bol, David K.
/ APPLICANT: Cardoni, Joan M.
/ APPLICANT: Rowley, Ronald B.
/ APPLICANT: Wong, Tai W.
/ APPLICANT: Lee, Francis Y.
/ TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
/ TITLE OF INVENTION: TYROSINE KINASE RECEPTORS
/ FILE REFERENCE: D0254 ND
/ CURRENT APPLICATION NUMBER: US/10/378.393
/ CURRENT FILING DATE: 2003-03-03
/ PRIOR APPLICATION NUMBER: US 60/360,889
/ PRIOR FILING DATE: 2002-03-01
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn Version 3.2
/ SEQ ID NO 11
/ LENGTH: 791
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-378-393-11

Query Match      100.0%; Score 54; DB 12; Length 791;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
Db 557 EBYLVPOQGF 566

RESULT 13
US-09-854-356-6
/ Sequence 6, Application US/09854356
/ Patent No. US2002017567A1
/ GENERAL INFORMATION:
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Gheysen, Dirk
/ APPLICANT: Corixa Corporation
/ APPLICANT: SmithKline Beecham Biologicals S. A.
/ TITLE OF INVENTION: HER-2/neu Fusion Proteins
/ FILE REFERENCE: 014058-009810PC
/ CURRENT APPLICATION NUMBER: US/09/854.356
/ CURRENT FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 09/493,480
/ PRIOR FILING DATE: 2000-01-28
/ PRIOR APPLICATION NUMBER: US 60/117,976
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 919
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:fusion protein
/ OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6

Query Match      100.0%; Score 54; DB 10; Length 919;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
Db 685 EBYLVPOQGF 694

RESULT 14
US-10-146-473-72
/ Sequence 72, Application US/10146473
/ Publication No. US2003010888A1
/ GENERAL INFORMATION:
/ APPLICANT: Scanlan, Matthew
```

```

/ APPLICANT: Gout, Ivan
/ APPLICANT: Stockert, Elisabeth
/ APPLICANT: Cure, Ali
/ APPLICANT: Chen, Yao-Tseng
/ APPLICANT: Old, Lloyd
/ TITLE OF INVENTION: Breast Cancer Antigens
/ FILE REFERENCE: L00461/70130 (JRV)
/ CURRENT APPLICATION NUMBER: US/10/146,473
/ PRIOR FILING DATE: 2002-05-15
/ PRIOR APPLICATION NUMBER: US 60/291,150
/ PRIOR FILING DATE: 2001-05-15
/ NUMBER OF SEQ ID NOS: 82
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 72
/ LENGTH: 1253
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-146-473-72

```

```

Query Match      100.0%; Score 54; DB 15; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      1 EEYLVPOQGF 10
        |||||
Db      1021 EEYLVPOQGF 1030

```

```

RESULT 15
US-09-811-123-9
/ Sequence 9, Application US/09811123
/ Patent No. US20020001587A1
/ GENERAL INFORMATION:
/ APPLICANT: Sharon Erickson
/ APPLICANT: Ralph Schwall
/ APPLICANT: Mark Sliwkowski
/ TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-EPFB
/ FILE REFERENCE: GENEENT.073A2
/ CURRENT APPLICATION NUMBER: US/09/811,123
/ CURRENT FILING DATE: 2001-03-16
/ PRIOR APPLICATION NUMBER: 60/238,327
/ PRIOR FILING DATE: 2000-10-05
/ PRIOR APPLICATION NUMBER: 09/602,530
/ PRIOR FILING DATE: 2000-06-23
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 9
/ LENGTH: 1255
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-811-123-9

```

```

Query Match      100.0%; Score 54; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      1 EEYLVPOQGF 10
        |||||
Db      1021 EEYLVPOQGF 1030

```

Search completed: February 7, 2004, 13:34:01
Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 16:16:50 ; Search time 45 Seconds
(without alignments)

5839.439 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815
Sequence: 1 MELALCRWGLLALLPGCA.....TRKGPYRNPYGLIDVPV 1255

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 04
Maximum Match 1004
Listing first 45 summaries

Database : Published Applications AA.*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	10 US-09-854-356-1	Sequence 1, Appl1
2	6815	100.0	1255	10 US-09-930-125-2	Sequence 2, Appl1
3	6815	100.0	1255	11 US-09-441-411-6	Sequence 6, Appl1
4	6815	100.0	1255	11 US-10-313-644-2	Sequence 2, Appl1
5	6815	100.0	1255	12 US-10-394-322A-17	Sequence 17, Appl1
6	6815	100.0	1255	12 US-10-245-871-553	Sequence 53, Appl1
7	6815	100.0	1255	15 US-10-207-655-45	Sequence 45, Appl1
8	6806	99.9	1255	9 US-09-811-123-9	Sequence 9, Appl1
9	6806	99.9	1255	9 US-09-811-115-3	Sequence 3, Appl1
10	6806	99.9	1255	12 US-10-207-498-6	Sequence 6, Appl1
11	6806	99.9	1255	12 US-10-338-730-2	Sequence 2, Appl1
12	6806	99.9	1255	12 US-10-322-892-4	Sequence 4, Appl1
13	6806	99.9	1255	12 US-10-272-437A-28	Sequence 28, Appl1
14	6806	99.9	1255	12 US-10-117-937-594	Sequence 594, Appl1
15	6806	99.9	1255	12 US-10-435-696-36	Sequence 36, Appl1

16	6806	99.9	1255	15 US-10-177-293-126	Sequence 126, Appl1
17	6804	99.8	1253	15 US-10-146-473-72	Sequence 72, Appl1
18	6803	99.8	1255	12 US-10-149-138-4641	Sequence 4641, Appl1
19	6757	99.1	1255	10 US-09-769-508-2	Sequence 2, Appl1
20	6001.5	88.1	1256	10 US-09-854-356-2	Sequence 2, Appl1
21	5997	88.0	1260	10 US-09-870-759-118	Sequence 118, Appl1
22	5997	88.0	1260	10 US-09-751-708A-118	Sequence 118, Appl1
23	5982.5	87.8	1256	10 US-09-854-356-14	Sequence 6, Appl1
24	4900	71.9	919	10 US-09-854-356-6	Sequence 6, Appl1
25	3798	55.7	690	12 US-10-412-804A-11	Sequence 11, Appl1
26	3776	55.4	712	10 US-09-854-356-7	Sequence 7, Appl1
27	3694.5	54.2	720	12 US-10-412-804A-6	Sequence 6, Appl1
28	3691	54.2	715	12 US-10-412-804A-10	Sequence 10, Appl1
29	3641	53.4	685	12 US-10-412-804A-4	Sequence 4, Appl1
30	3628	53.2	653	10 US-09-854-356-3	Sequence 3, Appl1
31	3590	52.7	645	9 US-09-921-161-1	Sequence 1, Appl1
32	3580	52.7	645	12 US-10-608-628-13	Sequence 13, Appl1
33	3590	52.7	645	15 US-10-268-501-13	Sequence 13, Appl1
34	3169	46.5	1210	9 US-09-725-433-2	Sequence 2, Appl1
35	3167	46.5	1210	12 US-10-394-322A-16	Sequence 16, Appl1
36	3110.5	45.6	654	10 US-09-854-356-8	Sequence 8, Appl1
37	3092	45.4	583	10 US-09-930-125-9	Sequence 9, Appl1
38	3092	45.4	587	10 US-09-930-125-8	Sequence 8, Appl1
39	3091	45.4	589	10 US-09-930-125-10	Sequence 10, Appl1
40	3091	45.4	600	10 US-09-930-125-11	Sequence 11, Appl1
41	3054	44.8	791	12 US-10-378-393-11	Sequence 11, Appl1
42	3040	44.6	570	12 US-10-378-393-15	Sequence 15, Appl1
43	3004.5	44.1	1308	12 US-09-940-101-2	Sequence 2, Appl1
44	3004.5	44.1	1308	12 US-10-394-322A-18	Sequence 18, Appl1
45	3004.5	44.1	1308	15 US-10-207-655-47	Sequence 47, Appl1

ALIGNMENTS

```
RESULT 1
US-09-854-356-1
Sequence 1, Application US/09854356
Patent No. US2002017567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Geyssen, Dirk
APPLICANT: Corixa Corporation
APPLICANT: SmithKline Beecham Biologicals S. A.
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356
CURRENT FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: US 09/493,480
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,976
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human HER-2/neu protein
NAME/KEY: DOMAIN
LOCATION: (1) (653)
OTHER INFORMATION: extracellular domain (ECD)
NAME/KEY: DOMAIN
LOCATION: (676) (1255)
OTHER INFORMATION: intracellular domain (ICD)
NAME/KEY: DOMAIN
LOCATION: (990) (1255)
OTHER INFORMATION: phosphorylation domain (PD)
NAME/KEY: DOMAIN
LOCATION: (990) (1048)
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
OTHER INFORMATION: portion (delta PD)
```

US-09-854-356-1

Query Match 100.0%; Score 6815; DB 10; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MELAALCRWGLLALPPGAASVQCTGTDMLRLPASPTHLDMLRHLYOGCQVQGNL 60
 1 MELAALCRWGLLALPPGAASVQCTGTDMLRLPASPTHLDMLRHLYOGCQVQGNL 60
 61 ELTYLPTNASTFLDIOEVOGYVLLAHNOVROVPLORLRIVGTOLFEEDNYALAVLDNG 120
 61 ELTYLPTNASTFLDIOEVOGYVLLAHNOVROVPLORLRIVGTOLFEEDNYALAVLDNG 120
 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMKDIFHKNOQLA 180
 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMKDIFHKNOQLA 180
 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLTRTVACGACARCKGRLPTDCCHEOC 240
 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLTRTVACGACARCKGRLPTDCCHEOC 240
 241 AAGCTGPKHSDCLACHFNHSGICEILCPALVTYNTDTPESMPNBSGRYTFGASCYTAAP 300
 241 AAGCTGPKHSDCLACHFNHSGICEILCPALVTYNTDTPESMPNBSGRYTFGASCYTAAP 300
 301 YNYLSTDVGSCTLVCPPLHNOEVTAAEDGTQRCCKSKPCAVCYGLGMEHLREVAVTSAN 360
 301 YNYLSTDVGSCTLVCPPLHNOEVTAAEDGTQRCCKSKPCAVCYGLGMEHLREVAVTSAN 360
 361 IOFPAGCKKIFGSLAPLPSFDGDDPASNTAPLOEOLQVETTEILGYLYISAMPDSL 420
 361 IOFPAGCKKIFGSLAPLPSFDGDDPASNTAPLOEOLQVETTEILGYLYISAMPDSL 420
 421 DLSVFNOLQVIRBRILHNGAYSLTLOGLISWGLRSLRELSGSLALIHNTHLCHFNTV 480
 421 DLSVFNOLQVIRBRILHNGAYSLTLOGLISWGLRSLRELSGSLALIHNTHLCHFNTV 480
 481 PMDOLFRNPHQALLHTANREDEBCVGBGLACHOLCARGHCMGPGPTQCVNCSQFLRGQEC 540
 481 PMDOLFRNPHQALLHTANREDEBCVGBGLACHOLCARGHCMGPGPTQCVNCSQFLRGQEC 540
 541 VESCRVLQGLPREVYANARHCLPCHPECOPONGSYTCGFPADQVACAHYKDPFCYARC 600
 541 VESCRVLQGLPREVYANARHCLPCHPECOPONGSYTCGFPADQVACAHYKDPFCYARC 600
 601 PSGVKEDLSVMPIMKPPDEEGACOPCPINCTHSCVDLDDGCPAEORASPLTSTISAVNG 660
 601 PSGVKEDLSVMPIMKPPDEEGACOPCPINCTHSCVDLDDGCPAEORASPLTSTISAVNG 660
 661 ILLVVLGVVFGILLIKRQOKIRKTYMRRLLOETELVEPLTPSGAMENQAMRLKETEL 720
 661 ILLVVLGVVFGILLIKRQOKIRKTYMRRLLOETELVEPLTPSGAMENQAMRLKETEL 720
 721 RKRYVLAGSAGFYVYKGIWIPDGENVKI PVAIKYLRKNTSPKANKELIDEAYVMAVGSP 780
 721 RKRYVLAGSAGFYVYKGIWIPDGENVKI PVAIKYLRKNTSPKANKELIDEAYVMAVGSP 780
 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHRENGRIGSODLWMCQIAGMSYLDVDR 840
 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHRENGRIGSODLWMCQIAGMSYLDVDR 840
 841 LVHRDLAARVLYKSPNHVKITDGLARLDDIDETEHADGKRPIMMALESILRRRFT 900
 841 LVHRDLAARVLYKSPNHVKITDGLARLDDIDETEHADGKRPIMMALESILRRRFT 900
 901 HOSDWSYGVTVWELMTFGAKPYDGIIPARBIPLDLKGERLPPOPICITIDVYIMVCKM 960
 901 HOSDWSYGVTVWELMTFGAKPYDGIIPARBIPLDLKGERLPPOPICITIDVYIMVCKM 960
 961 IDSECRAPRELVESEFRMADPPORPVVITQNEBIDGASPLDSTYRSLIBDDMGDVLDA 1020
 961 IDSECRAPRELVESEFRMADPPORPVVITQNEBIDGASPLDSTYRSLIBDDMGDVLDA 1020

RESULT 2

US-09-930-125-2

; Sequence 2, Application US/09930125
 ; Publication No. US20020193329A1
 ; GENERAL INFORMATION:

; APPLICANT: Hand-Zimmerman, Susan
 ; APPLICANT: Gheever, Martin A.
 ; APPLICANT: Roy, Teresa M.

; APPLICANT: Hodges, Michael J.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: McNeill, Patricia D.

; APPLICANT: Vedvick, Thomas S.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
 ; FILE REFERENCE: OF HER-2/NEU-ASSOCIATED MALIGNANCIES

; CURRENT APPLICATION NUMBER: US/09/930,125
 ; CURRENT FILING DATE: 2001-08-14
 ; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 1255
 ; TYPE: PR1

; ORGANISM: Homo sapien
 US-09-930-125-2

Query Match 100.0%; Score 6815; DB 10; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MELAALCRWGLLALPPGAASVQCTGTDMLRLPASPTHLDMLRHLYOGCQVQGNL 60
 1 MELAALCRWGLLALPPGAASVQCTGTDMLRLPASPTHLDMLRHLYOGCQVQGNL 60
 61 ELTYLPTNASTFLDIOEVOGYVLLAHNOVROVPLORLRIVGTOLFEEDNYALAVLDNG 120
 61 ELTYLPTNASTFLDIOEVOGYVLLAHNOVROVPLORLRIVGTOLFEEDNYALAVLDNG 120
 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMKDIFHKNOQLA 180
 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMKDIFHKNOQLA 180
 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLTRTVACGACARCKGRLPTDCCHEOC 240
 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLTRTVACGACARCKGRLPTDCCHEOC 240
 241 AAGCTGPKHSDCLACHFNHSGICEILCPALVTYNTDTPESMPNBSGRYTFGASCYTAAP 300
 241 AAGCTGPKHSDCLACHFNHSGICEILCPALVTYNTDTPESMPNBSGRYTFGASCYTAAP 300
 301 YNYLSTDVGSCTLVCPPLHNOEVTAAEDGTQRCCKSKPCAVCYGLGMEHLREVAVTSAN 360
 301 YNYLSTDVGSCTLVCPPLHNOEVTAAEDGTQRCCKSKPCAVCYGLGMEHLREVAVTSAN 360
 361 IOFPAGCKKIFGSLAPLPSFDGDDPASNTAPLOEOLQVETTEILGYLYISAMPDSL 420

```

Db      361 IOBPACCKIKFSGIAPLPSPPDDPASNTAPLOPEQLOVETLEITGYLISAMPDLSL 420
Qy      421 DLSVPONTQVIRGRILINNGAYSLTLOGISWLGRLSRLSGSLALIHNTLCPVHTV 480
Db      421 DLSVPONTQVIRGRILINNGAYSLTLOGISWLGRLSRLSGSLALIHNTLCPVHTV 480
Qy      481 PMDOLFRNPHOALHTNRPEDBCVGEGLACHOLCARGHGWGPGPTCCVNCOSQFLRQEC 540
Db      481 PMDOLFRNPHOALHTNRPEDBCVGEGLACHOLCARGHGWGPGPTCCVNCOSQFLRQEC 540
Qy      541 VECRCVLOGLPREYVNAHRLCPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Db      541 VECRCVLOGLPREYVNAHRLCPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Qy      601 PSQVKPDLSTYMPIMKPEDEGACQPCPINCTHSCVDLDKGCRAEORASPLTSIIISAVG 660
Db      601 PSQVKPDLSTYMPIMKPEDEGACQPCPINCTHSCVDLDKGCRAEORASPLTSIIISAVG 660
Qy      661 ILLVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKXTEL 720
Db      661 ILLVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKXTEL 720
Qy      721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVLEBNTSPKANKETLDEAYYMAVGSP 780
Db      721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVLEBNTSPKANKETLDEAYYMAVGSP 780
Qy      781 YVSRLLGICLTSTVOLVTQMLPYGCLLDHVRNRLGSDLLNMCQIAKMSYLEDR 840
Db      781 YVSRLLGICLTSTVOLVTQMLPYGCLLDHVRNRLGSDLLNMCQIAKMSYLEDR 840
Qy      841 LVHRDLAANVLYKSPNHVKITDGLARLLDIDETEHADGKVPKIMWALLESILRRFT 900
Db      841 LVHRDLAANVLYKSPNHVKITDGLARLLDIDETEHADGKVPKIMWALLESILRRFT 900
Qy      901 HOSDWSYGVTVWELMTGAKPYDGI PAREIPDLLEKGRLLPOPICTIDVYIMVYKCM 960
Db      901 HOSDWSYGVTVWELMTGAKPYDGI PAREIPDLLEKGRLLPOPICTIDVYIMVYKCM 960
Qy      961 IOBPACCKIKFSGIAPLPSPPDDPASNTAPLOPEQLOVETLEITGYLISAMPDLSL 420
Db      961 IOBPACCKIKFSGIAPLPSPPDDPASNTAPLOPEQLOVETLEITGYLISAMPDLSL 420
Qy      1021 BEYLVPQGFPCPDPAAGAGVNHHRSSSTRSGGDLTLGLEPSEBEPASPLAPSE 1080
Db      1021 BEYLVPQGFPCPDPAAGAGVNHHRSSSTRSGGDLTLGLEPSEBEPASPLAPSE 1080
Qy      1081 AGSDVFDGLGMAAGLQSLPTHDSPLQRYSEBTPVPLPSETDGYVAPLTCSPQPEYV 1140
Db      1081 AGSDVFDGLGMAAGLQSLPTHDSPLQRYSEBTPVPLPSETDGYVAPLTCSPQPEYV 1140
Qy      1141 NOPDVRPOPSPREGLPARPAGATLTERPKLSPKNGVYVDFPAGAVENPEYLTQ 1200
Db      1141 NOPDVRPOPSPREGLPARPAGATLTERPKLSPKNGVYVDFPAGAVENPEYLTQ 1200
Qy      1201 GGAAPQHPHPAPSPADNLVYWDQDPPEKGAPESTFKGPTAENBEYGLDVPV 1255
Db      1201 GGAAPQHPHPAPSPADNLVYWDQDPPEKGAPESTFKGPTAENBEYGLDVPV 1255

```

```

; SOFTWARE: FastSeq for windows version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-411-6

Query Match      100.0%; Score 6815; DB 11; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 METALCRMGILLALIPGAASQVCTGDMKRLRPSPTHLDMRLHYOGCGVOGNTL 60
Db      1 METALCRMGILLALIPGAASQVCTGDMKRLRPSPTHLDMRLHYOGCGVOGNTL 60
Qy      61 ETTYLPFNASLSPLOIQEVQGVLLAHNOVROPVLRLLIYNGTOLFEDNYALAVDNG 120
Db      61 ETTYLPFNASLSPLOIQEVQGVLLAHNOVROPVLRLLIYNGTOLFEDNYALAVDNG 120
Qy      121 DFLNNTTPVTCASPGGLREHQLRSLEILKGVLIQBNPOLCYODTILMKDIFPKNNOLA 180
Db      121 DFLNNTTPVTCASPGGLREHQLRSLEILKGVLIQBNPOLCYODTILMKDIFPKNNOLA 180
Qy      181 LTLIDNRSRACHPCSPMCKGSRCWGESSEDCSLTRTVAGGACARCKGPLPTDCHEOC 240
Db      181 LTLIDNRSRACHPCSPMCKGSRCWGESSEDCSLTRTVAGGACARCKGPLPTDCHEOC 240
Qy      241 AGCTGPKASDCLACHFNHSGICEIHCPLVYNTNTDPEFSMNPBGRYTFGASCVTACP 300
Db      241 AGCTGPKASDCLACHFNHSGICEIHCPLVYNTNTDPEFSMNPBGRYTFGASCVTACP 300
Qy      301 YVYLDVGSCTVCPILHNOVYTABDGTORCEKSCPKACAVCYGLGMEHLREVAATVSAN 360
Db      301 YVYLDVGSCTVCPILHNOVYTABDGTORCEKSCPKACAVCYGLGMEHLREVAATVSAN 360
Qy      361 IOBPACCKIKFSGIAPLPSPPDDPASNTAPLOPEQLOVETLEITGYLISAMPDLSL 420
Db      361 IOBPACCKIKFSGIAPLPSPPDDPASNTAPLOPEQLOVETLEITGYLISAMPDLSL 420
Qy      421 DLSVPONTQVIRGRILINNGAYSLTLOGISWLGRLSRLSGSLALIHNTLCPVHTV 480
Db      421 DLSVPONTQVIRGRILINNGAYSLTLOGISWLGRLSRLSGSLALIHNTLCPVHTV 480
Qy      481 PMDOLFRNPHOALHTNRPEDBCVGEGLACHOLCARGHGWGPGPTCCVNCOSQFLRQEC 540
Db      481 PMDOLFRNPHOALHTNRPEDBCVGEGLACHOLCARGHGWGPGPTCCVNCOSQFLRQEC 540
Qy      541 VECRCVLOGLPREYVNAHRLCPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Db      541 VECRCVLOGLPREYVNAHRLCPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Qy      601 PSQVKPDLSTYMPIMKPEDEGACQPCPINCTHSCVDLDKGCRAEORASPLTSIIISAVG 660
Db      601 PSQVKPDLSTYMPIMKPEDEGACQPCPINCTHSCVDLDKGCRAEORASPLTSIIISAVG 660
Qy      661 ILLVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKXTEL 720
Db      661 ILLVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKXTEL 720
Qy      721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVLEBNTSPKANKETLDEAYYMAVGSP 780
Db      721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVLEBNTSPKANKETLDEAYYMAVGSP 780
Qy      781 YVSRLLGICLTSTVOLVTQMLPYGCLLDHVRNRLGSDLLNMCQIAKMSYLEDR 840
Db      781 YVSRLLGICLTSTVOLVTQMLPYGCLLDHVRNRLGSDLLNMCQIAKMSYLEDR 840
Qy      841 LVHRDLAANVLYKSPNHVKITDGLARLLDIDETEHADGKVPKIMWALLESILRRFT 900
Db      841 LVHRDLAANVLYKSPNHVKITDGLARLLDIDETEHADGKVPKIMWALLESILRRFT 900
Qy      901 HOSDWSYGVTVWELMTGAKPYDGI PAREIPDLLEKGRLLPOPICTIDVYIMVYKCM 960
Db      901 HOSDWSYGVTVWELMTGAKPYDGI PAREIPDLLEKGRLLPOPICTIDVYIMVYKCM 960

```

```

RESULT 3
US-09-441-411-6
; Sequence 6, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Diels, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26

```

Db 901 HOSDWSYGVWELMTFGAKPYDGI PAREI PDLEKGERL POPPITIDVYIMVKCM 960
Qy 961 IDSECRPRFRELVSERFARMARDPQRFVYI QNEDLGPA SPUDSTFYRSILEDDBDMDGLVDA 1020
Db 961 IDSECRPRFRELVSERFARMARDPQRFVYI QNEDLGPA SPUDSTFYRSILEDDBDMDGLVDA 1020
Qy 1021 EBYLVPOQGFCCDPAPAGAGMYHHRSSSTSSGGSDLTGLSPSEEBAPRSLAPSEG 1080
Db 1021 EBYLVPOQGFCCDPAPAGAGMYHHRSSSTSSGGSDLTGLSPSEEBAPRSLAPSEG 1080
Qy 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPOPEYV 1140
Db 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPOPEYV 1140
Qy 1141 NOPDVPRPSPREGPLPARPAGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
Db 1141 NOPDVPRPSPREGPLPARPAGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPSPAFDNLVYWDODPBERGAPSTFKGTPTAENPEYLGIDVAV 1255
Db 1201 GGAAPQHPHPAPSPAFDNLVYWDODPBERGAPSTFKGTPTAENPEYLGIDVAV 1255

RESULT 4
US-10-313-644-2
Sequence 2, Application US/10313644
Publication No. US20030157119A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Cheever, Martin A.
APPLICANT: Hand-Zimmerman, Susan
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
TITLE OF INVENTION: AND VIRUS-ASSOCIATED MALIGNANCIES
FILE REFERENCE: 210121.483C3
CURRENT APPLICATION NUMBER: US/10/313.644
CURRENT FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1255
TYPE: PRN
ORGANISM: Homo. sapien
US-10-313-644-2

Query Match 100.0%; Score 6815; DB 12; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEALALCRWGLLALLPFGAASSTQVCTGDMKRLPASPETHLDMLEHLVQGCQVQGNL 60
Db 1 MEALALCRWGLLALLPFGAASSTQVCTGDMKRLPASPETHLDMLEHLVQGCQVQGNL 60
Qy 61 ELYTLPNABLSFLQDIQEVQGYVLIANOVQVPLQRLKRVKTOUPEFNVAALAVDNG 120
Db 61 ELYTLPNABLSFLQDIQEVQGYVLIANOVQVPLQRLKRVKTOUPEFNVAALAVDNG 120
Qy 121 DPLANTTPVYVAGSGGLRELOLRLSTLILKGGVLIQRPOLCYODTILMKDIFKKNOLA 180
Db 121 DPLANTTPVYVAGSGGLRELOLRLSTLILKGGVLIQRPOLCYODTILMKDIFKKNOLA 180
Qy 181 LTLIDTRSRACHPCSPWCKGSRGSSSBDQSLRTVACGAGCARKGPLPTDCHEQC 240
Db 181 LTLIDTRSRACHPCSPWCKGSRGSSSBDQSLRTVACGAGCARKGPLPTDCHEQC 240
Qy 241 AAGCTGKSHDCLALCPHNSGICELCPALVYNTOTPEBMPBERGRTYFGASCYVACP 300
Db 241 AAGCTGKSHDCLALCPHNSGICELCPALVYNTOTPEBMPBERGRTYFGASCYVACP 300
Qy 301 YNTLSTVGSCTVLCPLHNOEVTABDGTORCEKSKPCARCYGLGMEHLREVAAYTSAN 360
Db 301 YNTLSTVGSCTVLCPLHNOEVTABDGTORCEKSKPCARCYGLGMEHLREVAAYTSAN 360
Qy 361 IOBRAGCKKIFGLAFLPESFSDGPASNTAPLOEOLQVFEETLEITGYLYISAMPDLP 420

Db 361 IOBRAGCKKIFGLAFLPESFSDGPASNTAPLOEOLQVFEETLEITGYLYISAMPDLP 420
Qy 421 DLSVFQNLQVIRIRILHNGAYSLTLOGLISWIGLSRLRELSGALILHNTHLCFVHTV 480
Db 421 DLSVFQNLQVIRIRILHNGAYSLTLOGLISWIGLSRLRELSGALILHNTHLCFVHTV 480
Qy 481 PMDOLFNRNPHQALLHTANRPEDECVBEGLAHQLCARGHCWGGPTQCVNCSOFLRQEC 540
Db 481 PMDOLFNRNPHQALLHTANRPEDECVBEGLAHQLCARGHCWGGPTQCVNCSOFLRQEC 540
Qy 541 VEECRVLOGLPREYVNAHCLPCHRECOPONGSVTCGEPADCCVCAHKKDPFCVARC 600
Db 541 VEECRVLOGLPREYVNAHCLPCHRECOPONGSVTCGEPADCCVCAHKKDPFCVARC 600
Qy 601 PSGVKEPDLSTMPIMKPEDEGACQPCPINCTHSCVDLDDKGCFAEQASPLTSIIAIVG 660
Db 601 PSGVKEPDLSTMPIMKPEDEGACQPCPINCTHSCVDLDDKGCFAEQASPLTSIIAIVG 660
Qy 661 ILVVLVGVVFGILIKRQOKIRKYMRLLOETELVEPLTPSGAMNQOMRILKETEL 720
Db 661 ILVVLVGVVFGILIKRQOKIRKYMRLLOETELVEPLTPSGAMNQOMRILKETEL 720
Qy 721 RKRYUGSAGFVYVYGIWTPDEGNKIPVAILVLENTSPKANKEILDEAYVAVAGVSP 780
Db 721 RKRYUGSAGFVYVYGIWTPDEGNKIPVAILVLENTSPKANKEILDEAYVAVAGVSP 780
Qy 781 YVSRILGICLSTVQVLTQMLPYGCLLDHYENRGLSGODLNMCMQIAKMSYLDVR 840
Db 781 YVSRILGICLSTVQVLTQMLPYGCLLDHYENRGLSGODLNMCMQIAKMSYLDVR 840
Qy 841 LVHRLAARNVLVKSNHVKITDFGLARLIDIDETEHADGKVPIMWALLESILRRFT 900
Db 841 LVHRLAARNVLVKSNHVKITDFGLARLIDIDETEHADGKVPIMWALLESILRRFT 900
Qy 901 HOSDWSYGVWELMTFGAKPYDGI PAREI PDLEKGERL POPPITIDVYIMVKCM 960
Db 901 HOSDWSYGVWELMTFGAKPYDGI PAREI PDLEKGERL POPPITIDVYIMVKCM 960
Qy 961 IDSECRPRFRELVSERFARMARDPQRFVYI QNEDLGPA SPUDSTFYRSILEDDBDMDGLVDA 1020
Db 961 IDSECRPRFRELVSERFARMARDPQRFVYI QNEDLGPA SPUDSTFYRSILEDDBDMDGLVDA 1020
Qy 1021 EBYLVPOQGFCCDPAPAGAGMYHHRSSSTSSGGSDLTGLSPSEEBAPRSLAPSEG 1080
Db 1021 EBYLVPOQGFCCDPAPAGAGMYHHRSSSTSSGGSDLTGLSPSEEBAPRSLAPSEG 1080
Qy 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPOPEYV 1140
Db 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPOPEYV 1140
Qy 1141 NOPDVPRPSPREGPLPARPAGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
Db 1141 NOPDVPRPSPREGPLPARPAGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPSPAFDNLVYWDODPBERGAPSTFKGTPTAENPEYLGIDVAV 1255
Db 1201 GGAAPQHPHPAPSPAFDNLVYWDODPBERGAPSTFKGTPTAENPEYLGIDVAV 1255

RESULT 5
US-10-394-322A-17
Sequence 17, Application US/10394322A
Publication No. US20030232391A1
GENERAL INFORMATION:
APPLICANT: SENEZIS PHARMACEUTICALS, INC.
APPLICANT: Prescott, John C.
TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
FILE REFERENCE: 39750-0006 US
CURRENT APPLICATION NUMBER: US/10/394.322A
CURRENT FILING DATE: 2003-03-20
PRIOR APPLICATION NUMBER: US 60/366,892
PRIOR FILING DATE: 2002-03-21

NUMBER OF SEQ ID NOS: 70
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 17
 LENGTH: 1255
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-394-322A-17

Query Match
 Best Local Similarity 100.0%; Score 6815; DB 12; Length 1255;
 Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MELAALCRWGLIALLPFGAASSTQVCTGDMKRLPASPEHMLMHLVYGCQVVGNTL 60
 1 MELAALCRWGLIALLPFGAASSTQVCTGDMKRLPASPEHMLMHLVYGCQVVGNTL 60
 61 ELTYLPTNASLSFLQDIOEVGVYLIANOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
 61 ELTYLPTNASLSFLQDIOEVGVYLIANOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
 121 DELNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
 121 DELNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
 121 DELNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
 181 LTLIDNRSRACHPCSPCKSGSRGWSSSDCOSLTRTVCAAGCARCKGPLPTDCHEOC 240
 181 LTLIDNRSRACHPCSPCKSGSRGWSSSDCOSLTRTVCAAGCARCKGPLPTDCHEOC 240
 241 AAGCTGPKHSIDCLACLFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACP 300
 241 AAGCTGPKHSIDCLACLFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACP 300
 301 YNYLSTDVGSCTLYCPHLNOEVTADGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
 301 YNYLSTDVGSCTLYCPHLNOEVTADGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
 361 IOEPACKKIPGSLAPLPSFDDDPASNTAPLOPELOQVETLEITGYLYISAMPDLSL 420
 361 IOEPACKKIPGSLAPLPSFDDDPASNTAPLOPELOQVETLEITGYLYISAMPDLSL 420
 421 DLSVPONLQVIRGRILHNGAYSLTLQGLISWLGSLSLRELSGLALIHNNTHLCPVHTV 480
 421 DLSVPONLQVIRGRILHNGAYSLTLQGLISWLGSLSLRELSGLALIHNNTHLCPVHTV 480
 481 PWDQLFRNPHOALHTANRPEDECVSGEGLACHOLCARHGMGPGPTQCVNCSQPLRGQEC 540
 481 PWDQLFRNPHOALHTANRPEDECVSGEGLACHOLCARHGMGPGPTQCVNCSQPLRGQEC 540
 541 VEECRVLOGLPRRYNVAHACLPCHPCQOPONGSVTCFGEPADQCVACAHYKDPFCVARC 600
 541 VEECRVLOGLPRRYNVAHACLPCHPCQOPONGSVTCFGEPADQCVACAHYKDPFCVARC 600
 601 PEGVPRDLSYMPIMKPEDEGACOPPCINCHSCVLDLKGCPABORASPLTSIIISAVNG 660
 601 PEGVPRDLSYMPIMKPEDEGACOPPCINCHSCVLDLKGCPABORASPLTSIIISAVNG 660
 661 ILVVVLGVVFGILIKRQOKIRKYMRLLOETELVEBELTSPGAMPNOAKRIKETEL 720
 661 ILVVVLGVVFGILIKRQOKIRKYMRLLOETELVEBELTSPGAMPNOAKRIKETEL 720
 721 RRVKVLGSGAFGTIVKGIWIPGGENYKIVAIKVLRENTSPRANKIILDEAYMAGVSGP 780
 721 RRVKVLGSGAFGTIVKGIWIPGGENYKIVAIKVLRENTSPRANKIILDEAYMAGVSGP 780
 781 YVSRLLIGLCTSTVOLVTLMPYGLLDHVRNRGBLGSODLLNMCQIAKMSITLEDVR 840
 781 YVSRLLIGLCTSTVOLVTLMPYGLLDHVRNRGBLGSODLLNMCQIAKMSITLEDVR 840
 841 LVHRLDAAANVAVKSPNHVKTIDFGLARLLDIDETEHADGGKVPKMMALSSILRRRT 900
 841 LVHRLDAAANVAVKSPNHVKTIDFGLARLLDIDETEHADGGKVPKMMALSSILRRRT 900
 901 HOSDWSIGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRLLPQPICTIDVYMINVKCM 960

1201 HOSDWSIGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRLLPQPICTIDVYMINVKCM 960
 961 IDSECRPRPRELVESEFARMADPQRFVITQNEIDCGPASPLSTFRLSLLEDMDMDLVA 1020
 961 IDSECRPRPRELVESEFARMADPQRFVITQNEIDCGPASPLSTFRLSLLEDMDMDLVA 1020
 1021 EBYLVPQGFPCPDPAAGAGVHHRSSSTRSGGGLTLGLBPESEAPRSPAPSEG 1080
 1021 EBYLVPQGFPCPDPAAGAGVHHRSSSTRSGGGLTLGLBPESEAPRSPAPSEG 1080
 1081 AGSDVFPDGLGMAKGLQSLPTHDPSPLOXYSSDPVPLPSEFDGYVAPLTCSPQEVY 1140
 1081 AGSDVFPDGLGMAKGLQSLPTHDPSPLOXYSSDPVPLPSEFDGYVAPLTCSPQEVY 1140
 1141 NOPDVRPOPSPREGPLPAPRAGATLERPKTSLPGKNGVYKDYAFEGAVENPEYLTPO 1200
 1141 NOPDVRPOPSPREGPLPAPRAGATLERPKTSLPGKNGVYKDYAFEGAVENPEYLTPO 1200
 1201 GGAAPQHPAPPAPSPAFNDLYMDQDPERGAPSTFGPTANPEYLGADVPY 1255
 1201 GGAAPQHPAPPAPSPAFNDLYMDQDPERGAPSTFGPTANPEYLGADVPY 1255

RESULT 6
 US-10-245-871-553
 Sequence 553, Application US/10245871
 Publication No. US20030235594A1

GENERAL INFORMATION:
 APPLICANT: HUMPHREYS, ROBERT
 APPLICANT: X0, MINZHEN
 TITLE OF INVENTION: II-KEY/ANTIGENIC EPI TOPE HYBRID PEPTIDE VACCINES
 FILE REFERENCE: REH-2013
 CURRENT APPLICATION NUMBER: US/10/245, 871
 PRIOR FILING DATE: 2003-01-09
 PRIOR APPLICATION NUMBER: 10/197, 000
 PRIOR FILING DATE: 2002-07-17
 PRIOR APPLICATION NUMBER: 09/396, 813
 NUMBER OF SEQ ID NOS: 905
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 553
 LENGTH: 1255
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-245-871-553

Query Match
 Best Local Similarity 100.0%; Score 6815; DB 12; Length 1255;
 Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MELAALCRWGLIALLPFGAASSTQVCTGDMKRLPASPEHMLMHLVYGCQVVGNTL 60
 1 MELAALCRWGLIALLPFGAASSTQVCTGDMKRLPASPEHMLMHLVYGCQVVGNTL 60
 61 ELTYLPTNASLSFLQDIOEVGVYLIANOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
 61 ELTYLPTNASLSFLQDIOEVGVYLIANOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
 121 DELNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
 121 DELNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
 121 DELNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
 181 LTLIDNRSRACHPCSPCKSGSRGWSSSDCOSLTRTVCAAGCARCKGPLPTDCHEOC 240
 181 LTLIDNRSRACHPCSPCKSGSRGWSSSDCOSLTRTVCAAGCARCKGPLPTDCHEOC 240
 241 AAGCTGPKHSIDCLACLFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACP 300
 241 AAGCTGPKHSIDCLACLFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACP 300
 301 YNYLSTDVGSCTLYCPHLNOEVTADGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360

```

Db      301 YNLTSTDVSGCTLVCPCLHNOEVTAEADGTORCEKCSKPCARVCYGLGMEHLREVAVTSAN 360
Qy      361 IOEFACCKKI FGSIAFLPESFDSDPASNTAPLOPELOVETLEITGYLYISAMPDLSL 420
Db      361 IOEFACCKKI FGSIAFLPESFDSDPASNTAPLOPELOVETLEITGYLYISAMPDLSL 420
Qy      421 DLSVPONLOVIRGRILHNGAYSLTLQGLGISWLGSLSLRELSGLALIHNTHLCPVHTV 480
Db      421 DLSVPONLOVIRGRILHNGAYSLTLQGLGISWLGSLSLRELSGLALIHNTHLCPVHTV 480
Qy      481 PMOQLFRNPQALHHTANRPEDCEVGBGLACHOLCARGHCMGPGPTOCVNCOSQFLRGQEC 540
Db      481 PMOQLFRNPQALHHTANRPEDCEVGBGLACHOLCARGHCMGPGPTOCVNCOSQFLRGQEC 540
Qy      541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Db      541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Qy      601 PSGVKPDLSTMPYIMKPFDEBGACOPCPINCTHSCVDLDKGCPRAEORASPLTISIAYVG 660
Db      601 PSGVKPDLSTMPYIMKPFDEBGACOPCPINCTHSCVDLDKGCPRAEORASPLTISIAYVG 660
Qy      661 ILLVVLGVVFGILIKRQOKIRKRYTMRLLQETELVEPLTPSGAMNOAMRLKETEL 720
Db      661 ILLVVLGVVFGILIKRQOKIRKRYTMRLLQETELVEPLTPSGAMNOAMRLKETEL 720
Qy      721 RKRYVLGSAGFGTVYKGIWI PDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
Db      721 RKRYVLGSAGFGTVYKGIWI PDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
Qy      781 YSRRLGICLTSTVQVLTOLMPYGCCLDHVRENGRGLSGODLNMCMQIAKMSYLEDR 840
Db      781 YSRRLGICLTSTVQVLTOLMPYGCCLDHVRENGRGLSGODLNMCMQIAKMSYLEDR 840
Qy      841 LVHRDLAARVNLVSPNHVKITDGLARLLDIDETEHADGKVPKIMMALESILRRFT 900
Db      841 LVHRDLAARVNLVSPNHVKITDGLARLLDIDETEHADGKVPKIMMALESILRRFT 900
Qy      901 HOSDWSYGVYTWELMTFGAKFYDGI PABEIPDLLEKGERLPOPPICITIDVYIMVCKM 960
Db      901 HOSDWSYGVYTWELMTFGAKFYDGI PABEIPDLLEKGERLPOPPICITIDVYIMVCKM 960
Qy      961 IDSECRPRFELVSEFSRMAPDRFVYITONEDIGPSPLDSTYRSLIBDDMGDLVDA 1020
Db      961 IDSECRPRFELVSEFSRMAPDRFVYITONEDIGPSPLDSTYRSLIBDDMGDLVDA 1020
Qy      1021 BEYIVPOOGFPCPAPAGAGMVAHRRHSSSTRSGGDLTLGLPSEBEARSPPLASSEG 1080
Db      1021 BEYIVPOOGFPCPAPAGAGMVAHRRHSSSTRSGGDLTLGLPSEBEARSPPLASSEG 1080
Qy      1081 AGSDVPDGLCMGAAGKIQSLPTHDPSPLQRYSEDPVPLPSETDGVYAPVLTCSPOBEYV 1140
Db      1081 AGSDVPDGLCMGAAGKIQSLPTHDPSPLQRYSEDPVPLPSETDGVYAPVLTCSPOBEYV 1140
Qy      1141 NQDPVRPQPSRREGPLPAAPAPAGATLERPKTILSGKGVVYKDVAFAGGAVENBEYLTPQ 1200
Db      1141 NQDPVRPQPSRREGPLPAAPAPAGATLERPKTILSGKGVVYKDVAFAGGAVENBEYLTPQ 1200
Qy      1201 GGAAPOHPHPAPAPAPNLYYWDODPBERGAPSTFKGTPANPEYLGIADVP 1255
Db      1201 GGAAPOHPHPAPAPAPNLYYWDODPBERGAPSTFKGTPANPEYLGIADVP 1255

```

RESULT 7

```

US-10-207-655-45
; Sequence 45, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655

```

```

; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 45
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-45

Query Match      100.0%; Score 6815; DB 15; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MELAALCRGGLLALLPRAASVQVCTGDMRLRLPASERTHDMRLHYOGQOVQGNL 60
Db      1 MELAALCRGGLLALLPRAASVQVCTGDMRLRLPASERTHDMRLHYOGQOVQGNL 60
Qy      61 ELTYLPTNASLFLQDIQEVQGVYLIANQVQVPLQRLRIYRGTOLEFENVALAVLDNG 120
Db      61 ELTYLPTNASLFLQDIQEVQGVYLIANQVQVPLQRLRIYRGTOLEFENVALAVLDNG 120
Qy      121 DELNNTPTVYGASPGGLREIQRLSLTEILKGVLIQRPOLCYODITLMKDIFHKNOQA 180
Db      121 DELNNTPTVYGASPGGLREIQRLSLTEILKGVLIQRPOLCYODITLMKDIFHKNOQA 180
Qy      181 LTIIDTRSRACHPCSPCKGSRCKWGESSEDCSLTRTVACGACARCKGPLPDCCHEOC 240
Db      181 LTIIDTRSRACHPCSPCKGSRCKWGESSEDCSLTRTVACGACARCKGPLPDCCHEOC 240
Qy      241 AGCTGPKSDCLACAFHNSGICELHCPALVYNTDTEFSMNPBEGRYTFGASCVTACP 300
Db      241 AGCTGPKSDCLACAFHNSGICELHCPALVYNTDTEFSMNPBEGRYTFGASCVTACP 300
Qy      301 YNLTSTDVSGCTLVCPCLHNOEVTAEADGTORCEKCSKPCARVCYGLGMEHLREVAVTSAN 360
Db      301 YNLTSTDVSGCTLVCPCLHNOEVTAEADGTORCEKCSKPCARVCYGLGMEHLREVAVTSAN 360
Qy      361 IOEFACCKKI FGSIAFLPESFDSDPASNTAPLOPELOVETLEITGYLYISAMPDLSL 420
Db      361 IOEFACCKKI FGSIAFLPESFDSDPASNTAPLOPELOVETLEITGYLYISAMPDLSL 420
Qy      421 DLSVPONLOVIRGRILHNGAYSLTLQGLGISWLGSLSLRELSGLALIHNTHLCPVHTV 480
Db      421 DLSVPONLOVIRGRILHNGAYSLTLQGLGISWLGSLSLRELSGLALIHNTHLCPVHTV 480
Qy      481 PMOQLFRNPQALHHTANRPEDCEVGBGLACHOLCARGHCMGPGPTOCVNCOSQFLRGQEC 540
Db      481 PMOQLFRNPQALHHTANRPEDCEVGBGLACHOLCARGHCMGPGPTOCVNCOSQFLRGQEC 540
Qy      541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Db      541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Qy      601 PSGVKPDLSTMPYIMKPFDEBGACOPCPINCTHSCVDLDKGCPRAEORASPLTISIAYVG 660
Db      601 PSGVKPDLSTMPYIMKPFDEBGACOPCPINCTHSCVDLDKGCPRAEORASPLTISIAYVG 660
Qy      661 ILLVVLGVVFGILIKRQOKIRKRYTMRLLQETELVEPLTPSGAMNOAMRLKETEL 720
Db      661 ILLVVLGVVFGILIKRQOKIRKRYTMRLLQETELVEPLTPSGAMNOAMRLKETEL 720
Qy      721 RKRYVLGSAGFGTVYKGIWI PDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
Db      721 RKRYVLGSAGFGTVYKGIWI PDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
Qy      781 YSRRLGICLTSTVQVLTOLMPYGCCLDHVRENGRGLSGODLNMCMQIAKMSYLEDR 840
Db      781 YSRRLGICLTSTVQVLTOLMPYGCCLDHVRENGRGLSGODLNMCMQIAKMSYLEDR 840
Qy      841 LVHRDLAARVNLVSPNHVKITDGLARLLDIDETEHADGKVPKIMMALESILRRFT 900
Db      841 LVHRDLAARVNLVSPNHVKITDGLARLLDIDETEHADGKVPKIMMALESILRRFT 900

```

Qy 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREI PDLLEKGERL POPPICTIDVYMIWKCWM 960
 Db 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREI PDLLEKGERL POPPICTIDVYMIWKCWM 960
 Qy 961 IDSECRPRFRELVSFSSRMARDPQRFVVIQNEDELGPASPLDSTFYRSLBEDDMGDLVDA 1020
 Db 961 IDSECRPRFRELVSFSSRMARDPQRFVVIQNEDELGPASPLDSTFYRSLBEDDMGDLVDA 1020
 Qy 1021 BEVLVPOQGFPCDPAFGAGGVHHRSSSTRSGGDLTLGLPSEEEAPRSPPLAPSEG 1080
 Db 1021 BEVLVPOQGFPCDPAFGAGGVHHRSSSTRSGGDLTLGLPSEEEAPRSPPLAPSEG 1080
 Qy 1081 AGSDVDPDGLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVVALTCSPOPEYV 1140
 Db 1081 AGSDVDPDGLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVVALTCSPOPEYV 1140
 Qy 1141 NOPDVAPQPPSPREGLPAPAPAGATLERPKTSLPGKNGVAKDVFAGAVENPEYLTPO 1200
 Db 1141 NOPDVAPQPPSPREGLPAPAPAGATLERPKTSLPGKNGVAKDVFAGAVENPEYLTPO 1200
 Qy 1201 GGAAPQPPPPAPSPAFDNLVYWDODPPERGAPSTFGTPTAENPEYLGLDVAV 1255
 Db 1201 GGAAPQPPPPAPSPAFDNLVYWDODPPERGAPSTFGTPTAENPEYLGLDVAV 1255

RESULT 8

US-09-811-123-9
 / Sequence 9, Application US/09811123
 / Patent No. US20020001587A1
 / GENERAL INFORMATION:
 / APPLICANT: Sharon Erickson
 / APPLICANT: Ralph Schwall
 / APPLICANT: Mark Sliwowski
 / TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-EBB
 / TITLE OF INVENTION: ANTI-BODY-MAYTANSINOID CONJUGATES
 / FILE REFERENCE: GENENT. 073A2
 / CURRENT APPLICATION NUMBER: US/09/811,123
 / PRIOR FILING DATE: 2001-03-16
 / PRIOR APPLICATION NUMBER: 60/238,327
 / PRIOR FILING DATE: 2000-10-05
 / PRIOR APPLICATION NUMBER: 09/602,530
 / NUMBER OF SEQ ID NOS: 11
 / SOFTWARE: FaecSeq For Windows Version 4.0
 / SEQ ID NO 9
 / LENGTH: 1255
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-811-123-9

Query Match 99.9%; Score 6806; DB 9; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELALCCKWGLLALLPPASASTVCTGTDMKRLPASPETHLDMRLHYGCGVAVGNTL 60
 Db 1 MELALCCKWGLLALLPPASASTVCTGTDMKRLPASPETHLDMRLHYGCGVAVGNTL 60
 Qy 61 ELTYLPTNALSIFLQDIOEVGVYLAHNOVQVPLQRLRIYRGQLFEDNALAVLDNG 120
 Db 61 ELTYLPTNALSIFLQDIOEVGVYLAHNOVQVPLQRLRIYRGQLFEDNALAVLDNG 120
 Qy 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGVLLIQRNQOLCYOTIILMKDIFHKNOLA 180
 Db 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGVLLIQRNQOLCYOTIILMKDIFHKNOLA 180
 Qy 181 LTLIDTNSRACHPSCSPKCKSRGCSSEBDCOSLTRVTCAGACARCKGRLPTDCCHQC 240
 Db 181 LTLIDTNSRACHPSCSPKCKSRGCSSEBDCOSLTRVTCAGACARCKGRLPTDCCHQC 240
 Qy 241 AAGCTGPRKSDCLACLFHNSGICELHCPALVTNTDTPESMPNREGRTBASCTVACP 300
 Db 241 AAGCTGPRKSDCLACLFHNSGICELHCPALVTNTDTPESMPNREGRTBASCTVACP 300

Qy 301 YNLTSDVGSCTLVCPHLNQEVTAEADGTORCEKCSKPCARVCYGLGMEHLREVAATVSAN 360
 Db 301 YNLTSDVGSCTLVCPHLNQEVTAEADGTORCEKCSKPCARVCYGLGMEHLREVAATVSAN 360
 Qy 361 IQFACCKKIFGSLATLPSFDDPPLASNTAPLOPEOLQVETLEITGYLYISAMPDSL 420
 Db 361 IQFACCKKIFGSLATLPSFDDPPLASNTAPLOPEOLQVETLEITGYLYISAMPDSL 420
 Qy 421 DISFONLOVIRKRIILHNGAYSLTLQGLISWGLSLRELSGLALIHNTHLCPHTV 480
 Db 421 DISFONLOVIRKRIILHNGAYSLTLQGLISWGLSLRELSGLALIHNTHLCPHTV 480
 Qy 481 PMDOLFERNHQAHLHTANRPEDECVGEGLACHQLCARGHCWGPPTQCVNCSQFLRQEC 540
 Db 481 PMDOLFERNHQAHLHTANRPEDECVGEGLACHQLCARGHCWGPPTQCVNCSQFLRQEC 540
 Qy 541 VBERCVLQGLPRRYNARHCLCPHECOPONGSVTCFGBADCCVACAHYKDPFCVARC 600
 Db 541 VBERCVLQGLPRRYNARHCLCPHECOPONGSVTCFGBADCCVACAHYKDPFCVARC 600
 Qy 601 PSCVPRDLSYMPIMKRPDEEGACOPCINCTHSQVLDKGCAPRASPSTIVSAVG 660
 Db 601 PSCVPRDLSYMPIMKRPDEEGACOPCINCTHSQVLDKGCAPRASPSTIVSAVG 660
 Qy 661 ILVYVVLGVVFGILLKROOKIRKTYMRRLQETELVEPLTESGAMPNQAOMRILKETEL 720
 Db 661 ILVYVVLGVVFGILLKROOKIRKTYMRRLQETELVEPLTESGAMPNQAOMRILKETEL 720
 Qy 721 RKXKVLGSGAFGTIVKGIWIPGENYKIPVAIKVIRENTSPXANKIILDEAYYVAGVSP 780
 Db 721 RKXKVLGSGAFGTIVKGIWIPGENYKIPVAIKVIRENTSPXANKIILDEAYYVAGVSP 780
 Qy 781 YVSRLLIGLSTVOLVQMLMPYGLLDHVRNRGLSGODILNMCQIAKGSYLEDVR 840
 Db 781 YVSRLLIGLSTVOLVQMLMPYGLLDHVRNRGLSGODILNMCQIAKGSYLEDVR 840
 Qy 841 LVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETERYADGGKVPKXMALESILRRFT 900
 Db 841 LVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETERYADGGKVPKXMALESILRRFT 900
 Qy 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREI PDLLEKGERL POPPICTIDVYMIWKCWM 960
 Db 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREI PDLLEKGERL POPPICTIDVYMIWKCWM 960
 Qy 961 IDSECRPRFRELVSFSSRMARDPQRFVVIQNEDELGPASPLDSTFYRSLBEDDMGDLVDA 1020
 Db 961 IDSECRPRFRELVSFSSRMARDPQRFVVIQNEDELGPASPLDSTFYRSLBEDDMGDLVDA 1020
 Qy 1021 BEVLVPOQGFPCDPAFGAGGVHHRSSSTRSGGDLTLGLPSEEEAPRSPPLAPSEG 1080
 Db 1021 BEVLVPOQGFPCDPAFGAGGVHHRSSSTRSGGDLTLGLPSEEEAPRSPPLAPSEG 1080
 Qy 1081 AGSDVDPDGLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVVALTCSPOPEYV 1140
 Db 1081 AGSDVDPDGLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVVALTCSPOPEYV 1140
 Qy 1141 NOPDVAPQPPSPREGLPAPAPAGATLERPKTSLPGKNGVAKDVFAGAVENPEYLTPO 1200
 Db 1141 NOPDVAPQPPSPREGLPAPAPAGATLERPKTSLPGKNGVAKDVFAGAVENPEYLTPO 1200
 Qy 1201 GGAAPQPPPPAPSPAFDNLVYWDODPPERGAPSTFGTPTAENPEYLGLDVAV 1255
 Db 1201 GGAAPQPPPPAPSPAFDNLVYWDODPPERGAPSTFGTPTAENPEYLGLDVAV 1255

RESULT 9

US-09-811-115-3
 / Sequence 3, Application US/09811115
 / Patent No. US2002003573A1
 / GENERAL INFORMATION:
 / APPLICANT: Erickson, Sharon
 / APPLICANT: Schwall, Ralph

APPLICANT: King, Kathleen
TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
FILE REFERENCE: GENENT. 034A
CURRENT APPLICATION NUMBER: US/09/811,115
CURRENT FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/189,844
PRIOR FILING DATE: 2000-03-16
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 3
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-09-811-115-3

Query Match 99.9%; Score 6806; DB 9; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MEALACRMGLLALLPAGAASTOVCTGTMKRLPASPEHLDMLHLYOGCQVQGNL 60
1 MEALACRMGLLALLPAGAASTOVCTGTMKRLPASPEHLDMLHLYOGCQVQGNL 60
61 ELTYLPTNASTSLFLODIOEVQGYVLIHNOVROVPLQRLIRVGTQLFEDNYALAVLDNG 120
61 ELTYLPTNASTSLFLODIOEVQGYVLIHNOVROVPLQRLIRVGTQLFEDNYALAVLDNG 120
121 DPLANTPTVPGASPGGRLQLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
121 DPLANTPTVPGASPGGRLQLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
121 DPLANTPTVPGASPGGRLQLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
181 LTLIDNRSRACPCSPMCKGSRCKWSSSDCSLTFTVACGACGACGKPLPTDCCHBQC 240
181 LTLIDNRSRACPCSPMCKGSRCKWSSSDCSLTFTVACGACGACGKPLPTDCCHBQC 240
181 LTLIDNRSRACPCSPMCKGSRCKWSSSDCSLTFTVACGACGACGKPLPTDCCHBQC 240
241 AACCTGPKSHDCLACLFHNSGICELHCPALVTYNTDTPESMNPBGRYTFGASCTYACP 300
241 AACCTGPKSHDCLACLFHNSGICELHCPALVTYNTDTPESMNPBGRYTFGASCTYACP 300
241 AACCTGPKSHDCLACLFHNSGICELHCPALVTYNTDTPESMNPBGRYTFGASCTYACP 300
301 YNLTSDVSGCTVCPILHNOEVTADGTORCEKSKCAVCGGLGMEHLREYAVATSAN 360
301 YNLTSDVSGCTVCPILHNOEVTADGTORCEKSKCAVCGGLGMEHLREYAVATSAN 360
301 YNLTSDVSGCTVCPILHNOEVTADGTORCEKSKCAVCGGLGMEHLREYAVATSAN 360
361 IOEFAGCKIKFGLAPLPESFDGPASNTAPLOPEOLQVEETLEITGYLYISAMPDLP 420
361 IOEFAGCKIKFGLAPLPESFDGPASNTAPLOPEOLQVEETLEITGYLYISAMPDLP 420
361 IOEFAGCKIKFGLAPLPESFDGPASNTAPLOPEOLQVEETLEITGYLYISAMPDLP 420
421 DLSVPONLQVIRGRILHNGAVSLTLOGISWGLRSLRELISGLALIHNTHLCPVHTV 480
421 DLSVPONLQVIRGRILHNGAVSLTLOGISWGLRSLRELISGLALIHNTHLCPVHTV 480
421 DLSVPONLQVIRGRILHNGAVSLTLOGISWGLRSLRELISGLALIHNTHLCPVHTV 480
481 PMOOLPFPNPOALHTANRPDECEVGBGLACHQICAGHCHGPGPTCCVNCISOPLRQEC 540
481 PMOOLPFPNPOALHTANRPDECEVGBGLACHQICAGHCHGPGPTCCVNCISOPLRQEC 540
481 PMOOLPFPNPOALHTANRPDECEVGBGLACHQICAGHCHGPGPTCCVNCISOPLRQEC 540
541 VEECRVYQGLPREYVNAARHCLPCHPCOPONGSTCFEPZADQVACAHYDPPFCVARC 600
541 VEECRVYQGLPREYVNAARHCLPCHPCOPONGSTCFEPZADQVACAHYDPPFCVARC 600
541 VEECRVYQGLPREYVNAARHCLPCHPCOPONGSTCFEPZADQVACAHYDPPFCVARC 600
601 PSQVAPDLSYMPWPKPDEEGACOPCPINCTHSCVDLDDKGCAPBQASPLTISAVAG 660
601 PSQVAPDLSYMPWPKPDEEGACOPCPINCTHSCVDLDDKGCAPBQASPLTISAVAG 660
601 PSQVAPDLSYMPWPKPDEEGACOPCPINCTHSCVDLDDKGCAPBQASPLTISAVAG 660
661 ILVYVTVGVVFGILIKRQCKIRKTYMRLLQETELVEPLTPSGAMPQAOARILKETEL 720
661 ILVYVTVGVVFGILIKRQCKIRKTYMRLLQETELVEPLTPSGAMPQAOARILKETEL 720
661 ILVYVTVGVVFGILIKRQCKIRKTYMRLLQETELVEPLTPSGAMPQAOARILKETEL 720
721 RKVAVLSGAGTGYKGIWIPDGENVKI PVAIKYLRSTSPKANKELIDEAYVWAGVSP 780
721 RKVAVLSGAGTGYKGIWIPDGENVKI PVAIKYLRSTSPKANKELIDEAYVWAGVSP 780
721 RKVAVLSGAGTGYKGIWIPDGENVKI PVAIKYLRSTSPKANKELIDEAYVWAGVSP 780
781 YVSRLLGICLSTVQVLTQMPYGLLDHVENRGRIGSQDILNMCQIAKMSYLEDVR 840
781 YVSRLLGICLSTVQVLTQMPYGLLDHVENRGRIGSQDILNMCQIAKMSYLEDVR 840
781 YVSRLLGICLSTVQVLTQMPYGLLDHVENRGRIGSQDILNMCQIAKMSYLEDVR 840

781 YVSRLLGICLSTVQVLTQMPYGLLDHVENRGRIGSQDILNMCQIAKMSYLEDVR 840
841 LVHRDLAARNVYKSPNHVKITDFGLARLDDIDETERYHADGKVPKIMALSSILRRFT 900
841 LVHRDLAARNVYKSPNHVKITDFGLARLDDIDETERYHADGKVPKIMALSSILRRFT 900
841 LVHRDLAARNVYKSPNHVKITDFGLARLDDIDETERYHADGKVPKIMALSSILRRFT 900
901 HQSDVWSYGVTVWELMTFGAKYDGI PAREITDLEKGRLLPOPPCTIDVYIMKCMW 960
901 HQSDVWSYGVTVWELMTFGAKYDGI PAREITDLEKGRLLPOPPCTIDVYIMKCMW 960
901 HQSDVWSYGVTVWELMTFGAKYDGI PAREITDLEKGRLLPOPPCTIDVYIMKCMW 960
961 INSECRPRRELVSSESRARAPORFVITONEDLGPASPLDSFTYSLEDDMGDLVDA 1020
961 INSECRPRRELVSSESRARAPORFVITONEDLGPASPLDSFTYSLEDDMGDLVDA 1020
961 INSECRPRRELVSSESRARAPORFVITONEDLGPASPLDSFTYSLEDDMGDLVDA 1020
1021 EEYLVPOQGFPCDDPAPAGAGVHHRRSSSTRSGGDLTLGLEPSEEAAPSPLAPSEG 1080
1021 EEYLVPOQGFPCDDPAPAGAGVHHRRSSSTRSGGDLTLGLEPSEEAAPSPLAPSEG 1080
1021 EEYLVPOQGFPCDDPAPAGAGVHHRRSSSTRSGGDLTLGLEPSEEAAPSPLAPSEG 1080
1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLOKRSDEPTVPLPSETDGVAPLTCSPQREYV 1140
1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLOKRSDEPTVPLPSETDGVAPLTCSPQREYV 1140
1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLOKRSDEPTVPLPSETDGVAPLTCSPQREYV 1140
1141 NQPDVAPQPPSPREGPLPARPAGATLBRPKTLSPGKGVVQDVAFGAVENPEYLTPO 1200
1141 NQPDVAPQPPSPREGPLPARPAGATLBRPKTLSPGKGVVQDVAFGAVENPEYLTPO 1200
1141 NQPDVAPQPPSPREGPLPARPAGATLBRPKTLSPGKGVVQDVAFGAVENPEYLTPO 1200
1201 GGAAPQHPPEPAPAFDNLVYWDODPPERGAPSTFTKGTPTAENBEYGLDVPV 1255
1201 GGAAPQHPPEPAPAFDNLVYWDODPPERGAPSTFTKGTPTAENBEYGLDVPV 1255
1201 GGAAPQHPPEPAPAFDNLVYWDODPPERGAPSTFTKGTPTAENBEYGLDVPV 1255

RESULT 10

US-10-207-498-6
Sequence 6, Application US/10207498
Publication NO. US20030143568A1
GENERAL INFORMATION:
APPLICANT: Elizabeth Singer
APPLICANT: Kell Landgraf
APPLICANT: Dennis J. Slamon
APPLICANT: David Eisenberg
TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HERGULIN AND HER3
FILE REFERENCE: 30448:103-US-01
CURRENT APPLICATION NUMBER: US/10/207,498
CURRENT FILING DATE: 2002-07-29
PRIOR APPLICATION NUMBER: 60/308,431
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 6
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-10-207-498-6

Query Match 99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MEALACRMGLLALLPAGAASTOVCTGTMKRLPASPEHLDMLHLYOGCQVQGNL 60
1 MEALACRMGLLALLPAGAASTOVCTGTMKRLPASPEHLDMLHLYOGCQVQGNL 60
1 MEALACRMGLLALLPAGAASTOVCTGTMKRLPASPEHLDMLHLYOGCQVQGNL 60
61 ELTYLPTNASTSLFLODIOEVQGYVLIHNOVROVPLQRLIRVGTQLFEDNYALAVLDNG 120
61 ELTYLPTNASTSLFLODIOEVQGYVLIHNOVROVPLQRLIRVGTQLFEDNYALAVLDNG 120
61 ELTYLPTNASTSLFLODIOEVQGYVLIHNOVROVPLQRLIRVGTQLFEDNYALAVLDNG 120
121 DPLANTPTVPGASPGGRLQLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
121 DPLANTPTVPGASPGGRLQLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
121 DPLANTPTVPGASPGGRLQLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
181 LTLIDNRSRACPCSPMCKGSRCKWSSSDCSLTFTVACGACGACGKPLPTDCCHBQC 240
181 LTLIDNRSRACPCSPMCKGSRCKWSSSDCSLTFTVACGACGACGKPLPTDCCHBQC 240
181 LTLIDNRSRACPCSPMCKGSRCKWSSSDCSLTFTVACGACGACGKPLPTDCCHBQC 240

```

Db 181 LTLIDTNRBRACHPCSPKCKSGRCMBSSBDCQSLTRTVACGACGACRKGPLPTDCCHQC 240
Qy 241 AAGCTGPHNSCLACLPNHSIGICELHCPALVTNTDTFESMPNPEGRTYTGASCVTACP 300
Db 241 AAGCTGPHNSCLACLPNHSIGICELHCPALVTNTDTFESMPNPEGRTYTGASCVTACP 300
Qy 301 YNYLSTVGSCTLVCPHNOEVTABDGNORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
Db 301 YNYLSTVGSCTLVCPHNOEVTABDGNORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
Qy 361 IOBAGCKKIFGSLAPLPSFDPGASNTAPLOPEOLQVFTLEBTGYLISAMPDLP 420
Db 361 IOBAGCKKIFGSLAPLPSFDPGASNTAPLOPEOLQVFTLEBTGYLISAMPDLP 420
Qy 421 DLSVFNQLOVIRGRILHNGAYSLTLOGLISWGLRSLRELSGSLALIHNTHLCPVHTV 480
Db 421 DLSVFNQLOVIRGRILHNGAYSLTLOGLISWGLRSLRELSGSLALIHNTHLCPVHTV 480
Qy 481 PMDOLFRNPHOALHTANRPEDECVGEGLAGHOLCARGHCWGPGPTCCVNCOSQFLRQEC 540
Db 481 PMDOLFRNPHOALHTANRPEDECVGEGLAGHOLCARGHCWGPGPTCCVNCOSQFLRQEC 540
Qy 541 VESCRVLOGLPREYVNAHCLPCHECOPONGSVTCFGBADOCVAHAHYDPPCYARC 600
Db 541 VESCRVLOGLPREYVNAHCLPCHECOPONGSVTCFGBADOCVAHAHYDPPCYARC 600
Qy 601 PSQVPELSTYMPIMKFPDEBGAQPCPINCTHSCVDLDKGCAPABQASPLTSISAVNG 660
Db 601 PSQVPELSTYMPIMKFPDEBGAQPCPINCTHSCVDLDKGCAPABQASPLTSISAVNG 660
Qy 661 ILLVVLGVVFGILIKRROOKIRKTYMRLLQETELVEPLTPSGAMPNOQMILKETEL 720
Db 661 ILLVVLGVVFGILIKRROOKIRKTYMRLLQETELVEPLTPSGAMPNOQMILKETEL 720
Qy 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVALIKVLRNTSPKANKELLDAAVYMAVGSP 780
Db 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVALIKVLRNTSPKANKELLDAAVYMAVGSP 780
Qy 781 YVSRLLGICLTSTVQVLTQMPYGCCLLDHYENRNGRLGSDOLNMCQIAGKMSYLEDDR 840
Db 781 YVSRLLGICLTSTVQVLTQMPYGCCLLDHYENRNGRLGSDOLNMCQIAGKMSYLEDDR 840
Qy 841 LVHRDLAARVLYKSPHVKITPGLARLLIDBTEYHADGKVPKIMMALESILRRFT 900
Db 841 LVHRDLAARVLYKSPHVKITPGLARLLIDBTEYHADGKVPKIMMALESILRRFT 900
Qy 901 HOSDWSYGVTVLMTFGAKPYDGIIPAREIPDLLEKGERLQPPICCTIDVYIMKCM 960
Db 901 HOSDWSYGVTVLMTFGAKPYDGIIPAREIPDLLEKGERLQPPICCTIDVYIMKCM 960
Qy 961 IOBACRPRRELVSERBRMARDPQRFVIONEDLGPRASPLDSTFYSLLEDMDGLVDA 1020
Db 961 IOBACRPRRELVSERBRMARDPQRFVIONEDLGPRASPLDSTFYSLLEDMDGLVDA 1020
Qy 1021 BEYLVOQGFPCPDPAAGAGVYHHRSSSTSGGDLTLGLEPSEBEPASPLAPSEG 1080
Db 1021 BEYLVOQGFPCPDPAAGAGVYHHRSSSTSGGDLTLGLEPSEBEPASPLAPSEG 1080
Qy 1081 ACSDVDGDLGMAAGLOSLPTHDSPLORVSEDPVPLPSETDGVAPLTSPOPEYV 1140
Db 1081 ACSDVDGDLGMAAGLOSLPTHDSPLORVSEDPVPLPSETDGVAPLTSPOPEYV 1140
Qy 1141 NOPDVAPOPSPREGLPAPAPAGATLBRKTLSPKNGVYKDVPAFGAVENPELTPQ 1200
Db 1141 NOPDVAPOPSPREGLPAPAPAGATLBRKTLSPKNGVYKDVPAFGAVENPELTPQ 1200
Qy 1201 GGAAPPPHPPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 1255
Db 1201 GGAAPPPHPPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 1255

```

RESULT 11
US-10-338-730-2

```

; Sequence 2, Application US/10338730
; Publication No. US20030147905A1
; GENERAL INFORMATION:
; APPLICANT: Genzyme Corporation
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS
; FILE REFERENCE: 5017C
; CURRENT APPLICATION NUMBER: US/10/338, 730
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: US 09/527,487
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-338-730-2

```

```

Query Match 99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 METAALCRMGLLALLPPGAASQVCTGTDKRLPASPTHLDMLRHLVGGCOVQGNL 60
Db 1 METAALCRMGLLALLPPGAASQVCTGTDKRLPASPTHLDMLRHLVGGCOVQGNL 60
Qy 61 ELLYLPNLSLSTLOIOEVQGVYLAHNOVROPLORLAIVGTQLFEDNYALAVDNG 120
Db 61 ELLYLPNLSLSTLOIOEVQGVYLAHNOVROPLORLAIVGTQLFEDNYALAVDNG 120
Qy 121 DPLANTPTVYASPGGLRELQSLTEILKGVLIORNPOLCYODTILMDIIFHKNOA 180
Db 121 DPLANTPTVYASPGGLRELQSLTEILKGVLIORNPOLCYODTILMDIIFHKNOA 180
Qy 181 LTLIDTNRBRACHPCSPKCKSGRCMBSSBDCQSLTRTVACGACGACRKGPLPTDCCHQC 240
Db 181 LTLIDTNRBRACHPCSPKCKSGRCMBSSBDCQSLTRTVACGACGACRKGPLPTDCCHQC 240
Qy 241 AAGCTGPHNSCLACLPNHSIGICELHCPALVTNTDTFESMPNPEGRTYTGASCVTACP 300
Db 241 AAGCTGPHNSCLACLPNHSIGICELHCPALVTNTDTFESMPNPEGRTYTGASCVTACP 300
Qy 301 YNYLSTVGSCTLVCPHNOEVTABDGNORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
Db 301 YNYLSTVGSCTLVCPHNOEVTABDGNORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
Qy 361 IOBACRPRRELVSERBRMARDPQRFVIONEDLGPRASPLDSTFYSLLEDMDGLVDA 420
Db 361 IOBACRPRRELVSERBRMARDPQRFVIONEDLGPRASPLDSTFYSLLEDMDGLVDA 420
Qy 421 DLSVFNQLOVIRGRILHNGAYSLTLOGLISWGLRSLRELSGSLALIHNTHLCPVHTV 480
Db 421 DLSVFNQLOVIRGRILHNGAYSLTLOGLISWGLRSLRELSGSLALIHNTHLCPVHTV 480
Qy 481 PMDOLFRNPHOALHTANRPEDECVGEGLAGHOLCARGHCWGPGPTCCVNCOSQFLRQEC 540
Db 481 PMDOLFRNPHOALHTANRPEDECVGEGLAGHOLCARGHCWGPGPTCCVNCOSQFLRQEC 540
Qy 541 VESCRVLOGLPREYVNAHCLPCHECOPONGSVTCFGBADOCVAHAHYDPPCYARC 600
Db 541 VESCRVLOGLPREYVNAHCLPCHECOPONGSVTCFGBADOCVAHAHYDPPCYARC 600
Qy 601 PSQVPELSTYMPIMKFPDEBGAQPCPINCTHSCVDLDKGCAPABQASPLTSISAVNG 660
Db 601 PSQVPELSTYMPIMKFPDEBGAQPCPINCTHSCVDLDKGCAPABQASPLTSISAVNG 660
Qy 661 ILLVVLGVVFGILIKRROOKIRKTYMRLLQETELVEPLTPSGAMPNOQMILKETEL 720
Db 661 ILLVVLGVVFGILIKRROOKIRKTYMRLLQETELVEPLTPSGAMPNOQMILKETEL 720
Qy 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVALIKVLRNTSPKANKELLDAAVYMAVGSP 780

```

Db 721 RKXVIGSSAGFTVYKGIWIPDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAGVSP 780
QY 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYRENGRGLSODLNMCMQIAKMSYLEDDR 840
Db 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYRENGRGLSODLNMCMQIAKMSYLEDDR 840
QY 841 LVHRDLAARVNLVYKSPHVKITDFGLARLLDIDETEVHADGKVPKIMMALESTLRFRFT 900
Db 841 LVHRDLAARVNLVYKSPHVKITDFGLARLLDIDETEVHADGKVPKIMMALESTLRFRFT 900
QY 901 HQSDVMSYGTWELMTFGAKPYDGI PARETPDLEKGERLPORPCTIDVYMTVKCM 960
Db 901 HQSDVMSYGTWELMTFGAKPYDGI PARETPDLEKGERLPORPCTIDVYMTVKCM 960
QY 961 IDSECRPRFRELVESEFRMARDPQRFVVIQNEBGLSPASPLDSTFYRSLLEDDMDGLVDA 1020
Db 961 IDSECRPRFRELVESEFRMARDPQRFVVIQNEBGLSPASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 EBYLVPOQGFPCPDPAFGAGVMHRRSSSTRSGGDLTLGLEPSEBEAPRSLAPSEG 1080
Db 1021 EBYLVPOQGFPCPDPAFGAGVMHRRSSSTRSGGDLTLGLEPSEBEAPRSLAPSEG 1080
QY 1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLOKYSSEDPVPLPSETDGYVAPLTCSPQPEV 1140
Db 1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLOKYSSEDPVPLPSETDGYVAPLTCSPQPEV 1140
QY 1141 NOPDVRPQPSRPRGPRPAARPAATLERPKTISPGKNGVXVDFAFGAVENPEVLTPO 1200
Db 1141 NOPDVRPQPSRPRGPRPAARPAATLERPKTISPGKNGVXVDFAFGAVENPEVLTPO 1200
QY 1201 GGAAPQHPPPAPSPAFDNLVYMDODPPERGAPOSTFGTGTAEENPEVLTPO 1255
Db 1201 GGAAPQHPPPAPSPAFDNLVYMDODPPERGAPOSTFGTGTAEENPEVLTPO 1255

RESULT 12
US-10-322-892-4
/ Sequence 4, Application US/10322892
/ Publication No. US20030171257A1
/ GENERAL INFORMATION:
/ APPLICANT: STRBL, ROBERT C.
/ APPLICANT: SNEAD, MALCOLM L.
/ APPLICANT: XU, JIMMY
/ APPLICANT: VITETTA, EILEEN S.
/ APPLICANT: WILK, PETER J.
/ TITLE OF INVENTION: METHOD AND RELATED COMPOSITION EMPLOYING NANOSTRUCTURES
/ FILE REFERENCE: M07-505
/ CURRENT APPLICATION NUMBER: US/10/322, 892
/ PRIOR FILING DATE: 2002-12-18
/ PRIOR APPLICATION NUMBER: 60/342, 894
/ PRIOR FILING DATE: 2001-12-19
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: Patent Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 1255
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-322-892-4

Query Match 99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MELALCGMGLLALLPRGAASVCTGTDMKRLPASPETHLDMLRLHYGCGVQGNL 60
Db 1 MELALCGMGLLALLPRGAASVCTGTDMKRLPASPETHLDMLRLHYGCGVQGNL 60
QY 61 ELTYLPTNASLSPLODIOEVGVYLIANNOVROVPLQRLIVRGTOLEFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSPLODIOEVGVYLIANNOVROVPLQRLIVRGTOLEFEDNYALAVLDNG 120
QY 121 DPLANTTPTVAGSPGGLRELQSLTEILKGVLIQRPOLCYODTILMKOIFKHNNOLA 180
Db 121 DPLANTTPTVAGSPGGLRELQSLTEILKGVLIQRPOLCYODTILMKOIFKHNNOLA 180

Db 121 DPLANTTPTVAGSPGGLRELQSLTEILKGVLIQRPOLCYODTILMKOIFKHNNOLA 180
QY 181 LTIIDNRSRACHPCSPMCKGRSCEWSSSEDQSLTRTVACAGCARCKGPLPTDCHEQC 240
Db 181 LTIIDNRSRACHPCSPMCKGRSCEWSSSEDQSLTRTVACAGCARCKGPLPTDCHEQC 240
QY 241 AAGCTPKASDCLACHFNHSGICEHCPALVYNTDTEFSMNPBERYTFGASCYACP 300
Db 241 AAGCTPKASDCLACHFNHSGICEHCPALVYNTDTEFSMNPBERYTFGASCYACP 300
QY 301 YNLTSTDVSGCTLVCPFHNOEYTABDGTORCEKSKPCARVYGLGMEHREVAVTASAN 360
Db 301 YNLTSTDVSGCTLVCPFHNOEYTABDGTORCEKSKPCARVYGLGMEHREVAVTASAN 360
QY 361 IQSPACCKIKFGLAFIPESFDGDPASNTAPLOPBOLOVEETLEITGYLYISAMPDLP 420
Db 361 IQSPACCKIKFGLAFIPESFDGDPASNTAPLOPBOLOVEETLEITGYLYISAMPDLP 420
QY 421 DLSVFONLOYTRGRILLNNGAYSLTLOGLISWIGLSRLREGSLALIHNTHLCPVHTV 480
Db 421 DLSVFONLOYTRGRILLNNGAYSLTLOGLISWIGLSRLREGSLALIHNTHLCPVHTV 480
QY 481 PMPOLFPNPHQALLHTANRPEDECVGEGLAHQLCARHGMGPPTOCVNCQFLRGEC 540
Db 481 PMPOLFPNPHQALLHTANRPEDECVGEGLAHQLCARHGMGPPTOCVNCQFLRGEC 540
QY 541 VBCRCVLOGLPREYVNAHCLPCHBECOPONGSVTCFGEPAODCVACAHYKDPFCVARC 600
Db 541 VBCRCVLOGLPREYVNAHCLPCHBECOPONGSVTCFGEPAODCVACAHYKDPFCVARC 600
QY 601 PSQVKPDLSTMPYWKPPDEGACQPCINCTHSCVDLDDGCPAEQASPLTISIAYVG 660
Db 601 PSQVKPDLSTMPYWKPPDEGACQPCINCTHSCVDLDDGCPAEQASPLTISIAYVG 660
QY 661 ILLVVLGVYFGLIKRROOKIRKTYRRLLOETELVEPLTPSGAMENOQMRILKETEL 720
Db 661 ILLVVLGVYFGLIKRROOKIRKTYRRLLOETELVEPLTPSGAMENOQMRILKETEL 720
QY 721 RKXVIGSSAGFTVYKGIWIPDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAGVSP 780
Db 721 RKXVIGSSAGFTVYKGIWIPDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAGVSP 780
QY 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYRENGRGLSODLNMCMQIAKMSYLEDDR 840
Db 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYRENGRGLSODLNMCMQIAKMSYLEDDR 840
QY 841 LVHRDLAARVNLVYKSPHVKITDFGLARLLDIDETEVHADGKVPKIMMALESTLRFRFT 900
Db 841 LVHRDLAARVNLVYKSPHVKITDFGLARLLDIDETEVHADGKVPKIMMALESTLRFRFT 900
QY 901 HQSDVMSYGTWELMTFGAKPYDGI PARETPDLEKGERLPORPCTIDVYMTVKCM 960
Db 901 HQSDVMSYGTWELMTFGAKPYDGI PARETPDLEKGERLPORPCTIDVYMTVKCM 960
QY 961 IDSECRPRFRELVESEFRMARDPQRFVVIQNEBGLSPASPLDSTFYRSLLEDDMDGLVDA 1020
Db 961 IDSECRPRFRELVESEFRMARDPQRFVVIQNEBGLSPASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 EBYLVPOQGFPCPDPAFGAGVMHRRSSSTRSGGDLTLGLEPSEBEAPRSLAPSEG 1080
Db 1021 EBYLVPOQGFPCPDPAFGAGVMHRRSSSTRSGGDLTLGLEPSEBEAPRSLAPSEG 1080
QY 1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLOKYSSEDPVPLPSETDGYVAPLTCSPQPEV 1140
Db 1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLOKYSSEDPVPLPSETDGYVAPLTCSPQPEV 1140
QY 1141 NOPDVRPQPSRPRGPRPAARPAATLERPKTISPGKNGVXVDFAFGAVENPEVLTPO 1200
Db 1141 NOPDVRPQPSRPRGPRPAARPAATLERPKTISPGKNGVXVDFAFGAVENPEVLTPO 1200
QY 1201 GGAAPQHPPPAPSPAFDNLVYMDODPPERGAPOSTFGTGTAEENPEVLTPO 1255
Db 1201 GGAAPQHPPPAPSPAFDNLVYMDODPPERGAPOSTFGTGTAEENPEVLTPO 1255

RESULT 13
US-10-272-437A-28
Sequence 28, Application US/10272437A
Publication No. US20030216309A1
GENERAL INFORMATION:
APPLICANT: Krag, David N.
APPLICANT: Petro, Stephanie C.
APPLICANT: Oligino, Lyn
TITLE OF INVENTION: BINDING PEPTIDES SPECIFIC FOR THE EXTRACELLULAR DOMAIN OF ERBB2
TITLE OF INVENTION: US/10272437A
FILE REFERENCE: V00139, 70056, US
CURRENT APPLICATION NUMBER: US/10/272, 437A
PRIOR FILING DATE: 2002-10-15
PRIOR FILING DATE: 2001-10-12
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Patent version 3.1
SEQ ID NO 28
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-10-272-437A-28

Query Match 99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
1 METALALCKMGLLALPPGASTOVCTGMDKRLRASPETHLDMRLHYOGQVQVQNL 60
1 METALALCKMGLLALPPGASTOVCTGMDKRLRASPETHLDMRLHYOGQVQVQNL 60
61 ELTYLPTNASLFLDIOEVGVYLAHQVQVPLQRLIRYKGTOLFEDNALAVLDNG 120
61 ELTYLPTNASLFLDIOEVGVYLAHQVQVPLQRLIRYKGTOLFEDNALAVLDNG 120
121 DPLANTTPTVGTASPGGLRELQRLSLTEILKGVLIQRNPOLCYDITLWKDIFHKNQQA 180
121 DPLANTTPTVGTASPGGLRELQRLSLTEILKGVLIQRNPOLCYDITLWKDIFHKNQQA 180
121 DPLANTTPTVGTASPGGLRELQRLSLTEILKGVLIQRNPOLCYDITLWKDIFHKNQQA 180
181 LTLIDNRSRACHPCSPCKSGRSGESSEDCOSLTRTTCAGGACARCKPPLTDCHEQC 240
181 LTLIDNRSRACHPCSPCKSGRSGESSEDCOSLTRTTCAGGACARCKPPLTDCHEQC 240
181 LTLIDNRSRACHPCSPCKSGRSGESSEDCOSLTRTTCAGGACARCKPPLTDCHEQC 240
241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNDTFESMNPREGRTFGASCVTACP 300
241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNDTFESMNPREGRTFGASCVTACP 300
241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNDTFESMNPREGRTFGASCVTACP 300
301 YNYLSTDVGSCTLVCPHNOEVTABDGTQRCCKSPCARVCYGLAMEHLREVRATVSAN 360
301 YNYLSTDVGSCTLVCPHNOEVTABDGTQRCCKSPCARVCYGLAMEHLREVRATVSAN 360
301 YNYLSTDVGSCTLVCPHNOEVTABDGTQRCCKSPCARVCYGLAMEHLREVRATVSAN 360
361 IOEPAGCKKIRGSLAFLESPFGDPAANTAPLOPQLOVFEETLEBITGLYISAMPDILP 420
361 IOEPAGCKKIRGSLAFLESPFGDPAANTAPLOPQLOVFEETLEBITGLYISAMPDILP 420
361 IOEPAGCKKIRGSLAFLESPFGDPAANTAPLOPQLOVFEETLEBITGLYISAMPDILP 420
421 DLSVFNQLOVIRGRILHNGAVSLTQGLISWLGSLRELGSSGLTIHNNHLCPYTHV 480
421 DLSVFNQLOVIRGRILHNGAVSLTQGLISWLGSLRELGSSGLTIHNNHLCPYTHV 480
421 DLSVFNQLOVIRGRILHNGAVSLTQGLISWLGSLRELGSSGLTIHNNHLCPYTHV 480
481 PWDOLFRRPHOALHTANRPEDECVGEGSLAQHQLCARHCHWPGFTQCVNCSQFLRGEC 540
481 PWDOLFRRPHOALHTANRPEDECVGEGSLAQHQLCARHCHWPGFTQCVNCSQFLRGEC 540
541 VEECRVLOGLEFVYNAHCHLCPHCECOPONGSVTCFGBEADQVACAHYKDPFCVARC 600
541 VEECRVLOGLEFVYNAHCHLCPHCECOPONGSVTCFGBEADQVACAHYKDPFCVARC 600
601 PSQVAPDLSYMPIMKFPDEBAGACOPCPINCTHSQVLDLDDKGPABORASPLTSISAVVG 660
601 PSQVAPDLSYMPIMKFPDEBAGACOPCPINCTHSQVLDLDDKGPABORASPLTSISAVVG 660

661 ILVLVVLGVVFGILIRROOKIRKYMTRRLLOSTELVEPLTPSGAMPNOAMRLKETEL 720
661 ILVLVVLGVVFGILIRROOKIRKYMTRRLLOSTELVEPLTPSGAMPNOAMRLKETEL 720
721 RKYKVLGSAFGVYNGIWIPOGENKIPALIVALENTSPKANKELIDEAYVMAVGSP 780
721 RKYKVLGSAFGVYNGIWIPOGENKIPALIVALENTSPKANKELIDEAYVMAVGSP 780
781 YSRLLIGICLTSTVQVLTQIMPYGCLLDHVBENRGLSGODLNMCMQIAKGSYLEDVR 840
781 YSRLLIGICLTSTVQVLTQIMPYGCLLDHVBENRGLSGODLNMCMQIAKGSYLEDVR 840
841 LVHRDLAANVLYKSNHYKITDFGLARLLIDETEVHADGKVPKMALESILRRFT 900
841 LVHRDLAANVLYKSNHYKITDFGLARLLIDETEVHADGKVPKMALESILRRFT 900
901 HQSDWVSXGVTWELMTFAKPYDGIPIABEIPDLKGERLPQPICTIDVYIMVXCM 960
901 HQSDWVSXGVTWELMTFAKPYDGIPIABEIPDLKGERLPQPICTIDVYIMVXCM 960
961 IDSECRPRELVSERSRMAPDPORFVIONEDLGPASPLDSTFYRSLBEDDMDGLVDA 1020
961 IDSECRPRELVSERSRMAPDPORFVIONEDLGPASPLDSTFYRSLBEDDMDGLVDA 1020
1021 EBYLVPOQGFCCPDPAAGAGVHHRHRSSTRSGGDLTLGLEPSBEAPRSPAPSEG 1080
1021 EBYLVPOQGFCCPDPAAGAGVHHRHRSSTRSGGDLTLGLEPSBEAPRSPAPSEG 1080
1081 ACSVDFDGLGMAKGLQSLPTHDBSPLORYSEDPVLPSETDGVAPLTCSPQREYV 1140
1081 ACSVDFDGLGMAKGLQSLPTHDBSPLORYSEDPVLPSETDGVAPLTCSPQREYV 1140
1141 NOPDVAPOPSPREGPLPAARAGATLERPKLSPKSNVVDVAFGAVENPEYLTPO 1200
1141 NOPDVAPOPSPREGPLPAARAGATLERPKLSPKSNVVDVAFGAVENPEYLTPO 1200
1201 GGAPQPPHPPAPSAFNTLYTMDODPPERGAPSTFKTPTAENPEYLGADVPV 1255
1201 GGAPQPPHPPAPSAFNTLYTMDODPPERGAPSTFKTPTAENPEYLGADVPV 1255

RESULT 14
US-10-117-937-594
Sequence 594, Application US/10117937
Publication No. US20030220239A1
GENERAL INFORMATION:
APPLICANT: CTL IMMOBIO THERAPIES CORP.
APPLICANT: STIMARD, John, J.L.
APPLICANT: DIAMOND, David, C.
APPLICANT: Liu, Liding
APPLICANT: Xie, Zhidong
TITLE OF INVENTION: EPTOPE SEQUENCES
FILE REFERENCE: CTIIMM.027A
CURRENT APPLICATION NUMBER: US/10/117, 937
CURRENT FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/282, 211
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 60/337, 017
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/363, 210
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 602
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 594
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-10-117-937-594

Query Match 99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MELAALCRWGLIALLPAGAASQVCTGTDMLRLPASPTHLDMLRHLYOGCOVQGNL 60
 DB 1 MELAALCRWGLIALLPAGAASQVCTGTDMLRLPASPTHLDMLRHLYOGCOVQGNL 60
 QY 61 ELYTLPFNASLSFLQDIQEVQVYLIAHNOVROVPLQRLRIVRGTQLFEDNYALAVLDNG 120
 DB 61 ELYTLPFNASLSFLQDIQEVQVYLIAHNOVROVPLQRLRIVRGTQLFEDNYALAVLDNG 120
 QY 121 DPLANTTPVTGASPGGLRELQLSLREILKGVLIQNPOLCYODTILMKDI FHKNOOLA 180
 DB 121 DPLANTTPVTGASPGGLRELQLSLREILKGVLIQNPOLCYODTILMKDI FHKNOOLA 180
 QY 181 LTLIDNRSRACHPCSPMCKSGSRCKWGBSSBDQSLTFRTVACGACRCKGPLPTDCCHBQC 240
 DB 181 LTLIDNRSRACHPCSPMCKSGSRCKWGBSSBDQSLTFRTVACGACRCKGPLPTDCCHBQC 240
 QY 241 MAGCTPRKSDCLACLFHNSGICELHCPALVYNTDTPFSMPNBERYTFGASCVTACP 300
 DB 241 MAGCTPRKSDCLACLFHNSGICELHCPALVYNTDTPFSMPNBERYTFGASCVTACP 300
 QY 301 YNLTSTDVGSCTVCPPLHNOEVTABDGTORCEKSKRCARVCYGLGMEHLREYRAVTSAN 360
 DB 301 YNLTSTDVGSCTVCPPLHNOEVTABDGTORCEKSKRCARVCYGLGMEHLREYRAVTSAN 360
 QY 361 IQBFACCKKIFGSLAFLPESFDDGPASNTAPLOPEOLQVETLEITGYLYISAMPDSL 420
 DB 361 IQBFACCKKIFGSLAFLPESFDDGPASNTAPLOPEOLQVETLEITGYLYISAMPDSL 420
 QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRLSRLSGSLALIHNTHLCFVHTV 480
 DB 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRLSRLSGSLALIHNTHLCFVHTV 480
 QY 481 PMDOLFPNPHQALHTANRPEDBCVGBGLACHOLCARGHCWGPPTQVNCOSQFLRGQEC 540
 DB 481 PMDOLFPNPHQALHTANRPEDBCVGBGLACHOLCARGHCWGPPTQVNCOSQFLRGQEC 540
 QY 541 VEECRVLQGLPREVYNAHCLPCHBECQOPONGSVTCFPEADQCVACAHYKDPFCVAKC 600
 DB 541 VEECRVLQGLPREVYNAHCLPCHBECQOPONGSVTCFPEADQCVACAHYKDPFCVAKC 600
 QY 601 PEGVYKPLSTMPYKPFDEBGAQCPICNCHSCVDLDDKCPABQASPLTSLVSAVVG 660
 DB 601 PEGVYKPLSTMPYKPFDEBGAQCPICNCHSCVDLDDKCPABQASPLTSLVSAVVG 660
 QY 661 ILLVVLGVVFGILIKRQOKIKRYTMRRLQETELVEPLTPSGAMPNOAMRLIKETEL 720
 DB 661 ILLVVLGVVFGILIKRQOKIKRYTMRRLQETELVEPLTPSGAMPNOAMRLIKETEL 720
 QY 721 RKYKVLGSGAFGTVYKGIWIPDGENVKI PVAIKVLRNTSPKANKELIDEAYVMAVGGSP 780
 DB 721 RKYKVLGSGAFGTVYKGIWIPDGENVKI PVAIKVLRNTSPKANKELIDEAYVMAVGGSP 780
 QY 781 YVSRLLGICLTSTQVLTQMLPYGCLLDHRENRGRGLSODLLMWCMQIAGMSYLIBDVR 840
 DB 781 YVSRLLGICLTSTQVLTQMLPYGCLLDHRENRGRGLSODLLMWCMQIAGMSYLIBDVR 840
 QY 841 LVHRDLAABVNLVYKSPHNVKITDFFGLARLLDIDETBYHADGKYPIMMALESILRRFT 900
 DB 841 LVHRDLAABVNLVYKSPHNVKITDFFGLARLLDIDETBYHADGKYPIMMALESILRRFT 900
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPPICITIDVYMIWVKCM 960
 DB 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPPICITIDVYMIWVKCM 960
 QY 961 IDSECRPRFRELVEBFSHMAADPOQFVYIQNEBDLGPASPLDSTFRSLLEDDBDGDVDA 1020
 DB 961 IDSECRPRFRELVEBFSHMAADPOQFVYIQNEBDLGPASPLDSTFRSLLEDDBDGDVDA 1020
 QY 1021 BEYLVPOGFPCCPDAPAGAGVHHRHSSTRSGGDLTGLBSESEEARSPAPASEG 1080
 DB 1021 BEYLVPOGFPCCPDAPAGAGVHHRHSSTRSGGDLTGLBSESEEARSPAPASEG 1080
 QY 1081 AGSDVFPDGLMGAAKGLQSLPTHDPSPLORSYSDPTVLPSETDGVYAPLTCSBPQBYV 1140

DB 1081 AGSDVFPDGLMGAAKGLQSLPTHDPSPLORSYSDPTVLPSETDGVYAPLTCSBPQBYV 1140
 QY 1141 NQPDVAPQPPSPREBGLPAARAGATLERPKTLPKNGVYKDVAFGAVENPEYLTPO 1200
 DB 1141 NQPDVAPQPPSPREBGLPAARAGATLERPKTLPKNGVYKDVAFGAVENPEYLTPO 1200
 QY 1201 GGAAPQHPHPPAFSPAFDNLVYWDODPPERGAPSTFKGTPTAENBPLYGLDVPV 1255
 DB 1201 GGAAPQHPHPPAFSPAFDNLVYWDODPPERGAPSTFKGTPTAENBPLYGLDVPV 1255
 RESULT 15
 US-10-435-696-36
 ; Sequence 36, Application US/10435696
 ; Publication NO. US20040018525A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wirtz, Ralph
 ; APPLICANT: Nunes, Marc
 ; APPLICANT: Kallabie, Harald
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS
 ; FILE REFERENCE: Lea 36 108
 ; CURRENT APPLICATION NUMBER: US/10/435,696
 ; CURRENT FILING DATE: 2003-05-09
 ; PRIOR APPLICATION NUMBER: EP03003112.4
 ; PRIOR FILING DATE: 2003-02-13
 ; PRIOR APPLICATION NUMBER: EP02010291.9
 ; NUMBER OF SEQ ID NOS: 314
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 36
 ; LENGTH: 1255
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-435-696-36
 Query Match 99.9%; Score 6806; DB 12; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MELAALCRWGLIALLPAGAASQVCTGTDMLRLPASPTHLDMLRHLYOGCOVQGNL 60
 DB 1 MELAALCRWGLIALLPAGAASQVCTGTDMLRLPASPTHLDMLRHLYOGCOVQGNL 60
 QY 61 ELYTLPFNASLSFLQDIQEVQVYLIAHNOVROVPLQRLRIVRGTQLFEDNYALAVLDNG 120
 DB 61 ELYTLPFNASLSFLQDIQEVQVYLIAHNOVROVPLQRLRIVRGTQLFEDNYALAVLDNG 120
 QY 121 DPLANTTPVTGASPGGLRELQLSLREILKGVLIQNPOLCYODTILMKDI FHKNOOLA 180
 DB 121 DPLANTTPVTGASPGGLRELQLSLREILKGVLIQNPOLCYODTILMKDI FHKNOOLA 180
 QY 181 LTLIDNRSRACHPCSPMCKSGSRCKWGBSSBDQSLTFRTVACGACRCKGPLPTDCCHBQC 240
 DB 181 LTLIDNRSRACHPCSPMCKSGSRCKWGBSSBDQSLTFRTVACGACRCKGPLPTDCCHBQC 240
 QY 241 MAGCTPRKSDCLACLFHNSGICELHCPALVYNTDTPFSMPNBERYTFGASCVTACP 300
 DB 241 MAGCTPRKSDCLACLFHNSGICELHCPALVYNTDTPFSMPNBERYTFGASCVTACP 300
 QY 301 YNLTSTDVGSCTVCPPLHNOEVTABDGTORCEKSKRCARVCYGLGMEHLREYRAVTSAN 360
 DB 301 YNLTSTDVGSCTVCPPLHNOEVTABDGTORCEKSKRCARVCYGLGMEHLREYRAVTSAN 360
 QY 361 IQBFACCKKIFGSLAFLPESFDDGPASNTAPLOPEOLQVETLEITGYLYISAMPDSL 420
 DB 361 IQBFACCKKIFGSLAFLPESFDDGPASNTAPLOPEOLQVETLEITGYLYISAMPDSL 420
 QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRLSRLSGSLALIHNTHLCFVHTV 480
 DB 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRLSRLSGSLALIHNTHLCFVHTV 480

QY 481 PWDOLFRNPHQALLHTANRPEDECVGEGLACHOLCARGHCMGPRPTQCVNCSQFLRGQEC 540
DB 481 PWDOLFRNPHQALLHTANRPEDECVGEGLACHOLCARGHCMGPRPTQCVNCSQFLRGQEC 540
QY 541 VEECCVLOGLPREYVYNAHCLPCHPECOPONGSVTCFEPADOCVACAHYKDPFCVAC 600
DB 541 VEECCVLOGLPREYVYNAHCLPCHPECOPONGSVTCFEPADOCVACAHYKDPFCVAC 600
QY 601 PSGVXPDLSPYMPIMKRPDEEGACQPCPINCTHSCVDLDDKGCBAQORASPLTSIIISAVVG 660
DB 601 PSGVXPDLSPYMPIMKRPDEEGACQPCPINCTHSCVDLDDKGCBAQORASPLTSIIISAVVG 660
QY 661 ILLVVLGVVFGILLIKRROOKIRKTYMRLLQETELVEPLTPSGAMPNOAMRILKETEL 720
DB 661 ILLVVLGVVFGILLIKRROOKIRKTYMRLLQETELVEPLTPSGAMPNOAMRILKETEL 720
QY 721 RKVKYLSGSGAFSTVYKGIWIPDGENVKIPVAIKVLENTSPKANKKILDEAYMAGVSGP 780
DB 721 RKVKYLSGSGAFSTVYKGIWIPDGENVKIPVAIKVLENTSPKANKKILDEAYMAGVSGP 780
QY 781 YVSRLLGICLTSTVQLVTLQMPYGLLDHVRNENRGLSGODLLNMCQIAKMSYLEDV 840
DB 781 YVSRLLGICLTSTVQLVTLQMPYGLLDHVRNENRGLSGODLLNMCQIAKMSYLEDV 840
QY 841 LVHRDLAARNVYKSPNHVKITDFGLARLLDIDETRYHADGKVPKXMALESILRRFT 900
DB 841 LVHRDLAARNVYKSPNHVKITDFGLARLLDIDETRYHADGKVPKXMALESILRRFT 900
QY 901 HOSDVMYGVTVWMLTFCAPYDGI BAREIPDLKEXERLPOPICTIDVYIMVCMW 960
DB 901 HOSDVMYGVTVWMLTFCAPYDGI BAREIPDLKEXERLPOPICTIDVYIMVCMW 960
QY 961 IDSECRPRRELVSERSMARDPORFVVIQNEIDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRRELVSERSMARDPORFVVIQNEIDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 BEYLVPOCGFCPPDPAFGAGVHHRHRSSTRSGGDLTLGLEPSEBEAPRSPLAPSEG 1080
DB 1021 BEYLVPOCGFCPPDPAFGAGVHHRHRSSTRSGGDLTLGLEPSEBEAPRSPLAPSEG 1080
QY 1081 AGSDVFDGDLGAGAKGLQSLPTHDPSPLOXYSEDPVPLPSETDGYVAPLTCSPQPEV 1140
DB 1081 AGSDVFDGDLGAGAKGLQSLPTHDPSPLOXYSEDPVPLPSETDGYVAPLTCSPQPEV 1140
QY 1141 NOPDVRPOPSPREBEPPLPAARPAATLBRPXTLSPKNGVNVKOVFAFGAVENPEYLTPQ 1200
DB 1141 NOPDVRPOPSPREBEPPLPAARPAATLBRPXTLSPKNGVNVKOVFAFGAVENPEYLTPQ 1200
QY 1201 GGAAPQPHPPPAFPAFNDLYYWDQPPRGAAPSTFKGPTAENPEYLGLDVVP 1255
DB 1201 GGAAPQPHPPPAFPAFNDLYYWDQPPRGAAPSTFKGPTAENPEYLGLDVVP 1255

Search completed: February 9, 2004, 16:22:57
Job time : 50 secs

XX MPI; 1996-455361/45.
 DR N-PSDB; AAT40739.
 XX DNA encoding HER-2-neu poly:peptide(s) - used for prevention or
 PT treatment of malignancies with which the HER-2/neu oncogene is
 PT associated
 XX
 PS Claim 2; Page 56-61; 71pp; English.
 XX
 CC Human HER-2/neu protein (AAM0111), also called p185 or c-erbB2, is
 CC the product of the HER-2/neu oncogene (see also AAT40739). The
 CC protein is over-expressed in various cancers, including breast,
 CC ovarian, colon, lung and prostate. The intracellular domain of the
 CC protein can be used to immunise an animal against a malignancy with
 CC which the oncogene is associated. The polypeptide can be produced
 CC in transformed host cells for use in immunisation. Alternatively,
 CC animal cells are transfected in vivo or ex vivo with a viral vector
 CC that directs expression of the polypeptide.
 CC
 SQ Sequence 1255 AA;
 Query Match 100.0%; Score 6815; DB 17; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MELAALCRNGLLALALPGAASTQVCTGDMKRLPASPEHLDMRLHYQGQVQGNL 60
 1 MELAALCRNGLLALALPGAASTQVCTGDMKRLPASPEHLDMRLHYQGQVQGNL 60
 61 ELTYLPTNALSLSLQDIQEVQGYVLAHNOVROVPLQRLIRVGTQLFEDNVALAVLDNG 120
 61 ELTYLPTNALSLSLQDIQEVQGYVLAHNOVROVPLQRLIRVGTQLFEDNVALAVLDNG 120
 121 DPLANTPTVTSAPGRLRELQASLTREILKGVLIQNPOLCYOTILMDITRKNQOLA 180
 121 DPLANTPTVTSAPGRLRELQASLTREILKGVLIQNPOLCYOTILMDITRKNQOLA 180
 181 LTLIDTRSRACHPGSPKMGSRKMGSSSEDCSLTRTVAGGACARCKGPLPTDCHEOC 240
 181 LTLIDTRSRACHPGSPKMGSRKMGSSSEDCSLTRTVAGGACARCKGPLPTDCHEOC 240
 241 AAGCTGPKISDCLACIHFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCYAC 300
 241 AAGCTGPKISDCLACIHFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCYAC 300
 301 YNYLSTDVSGCTVCPHNOEYVABDGTORCKSCSKPCAVCYGLGMEHLRVRATVSAN 360
 301 YNYLSTDVSGCTVCPHNOEYVABDGTORCKSCSKPCAVCYGLGMEHLRVRATVSAN 360
 361 IOEPACCKKIFGSLAFLPESFPDGPASNTAPLOPQLOVETBEITGYLYISAMPDLP 420
 361 IOEPACCKKIFGSLAFLPESFPDGPASNTAPLOPQLOVETBEITGYLYISAMPDLP 420
 421 DLSVFQNLQVIRGRIILHNGAYSLTLOGLISWLGSLSLRELSSGALIIHNTHLCPVHTV 480
 421 DLSVFQNLQVIRGRIILHNGAYSLTLOGLISWLGSLSLRELSSGALIIHNTHLCPVHTV 480
 481 PMDOLFRNPHQALLHFRANREDECVGEGALCHQLCARHGWGSGPTQCVNCSQPLAGOC 540
 481 PMDOLFRNPHQALLHFRANREDECVGEGALCHQLCARHGWGSGPTQCVNCSQPLAGOC 540
 541 VEECRVLOGLPREYVNAHCLPCHPECOFNGSVTCGFPADOCVACAHXDPFCVARC 600
 541 VEECRVLOGLPREYVNAHCLPCHPECOFNGSVTCGFPADOCVACAHXDPFCVARC 600
 601 PSGVKDLSMPYKPFDEEGACOPCEINCTHS CVLDLDKGCAPAEQASPLTISAVVG 660
 601 PSGVKDLSMPYKPFDEEGACOPCEINCTHS CVLDLDKGCAPAEQASPLTISAVVG 660
 661 ILLVVLGVVFGILLIRROOKIRKTYMRRLLOSTELVEPLTSGGAMPNOQRIKXETEL 720
 661 ILLVVLGVVFGILLIRROOKIRKTYMRRLLOSTELVEPLTSGGAMPNOQRIKXETEL 720

QY 721 RAYKVLGSGAFGVYKGIWIPDGENYKIPVAIKVLEBNTSPKANKELDEAYVAVAGVSP 780
 DB 721 RAYKVLGSGAFGVYKGIWIPDGENYKIPVAIKVLEBNTSPKANKELDEAYVAVAGVSP 780
 QY 781 YNSRLIGICLTSVQVLTQMLPFGCLLDHVRERKGLSGSDLLNWCQIAKGSYLEDYR 840
 DB 781 YNSRLIGICLTSVQVLTQMLPFGCLLDHVRERKGLSGSDLLNWCQIAKGSYLEDYR 840
 QY 841 LVHRDLAARNVLYKSNHYKITDPFGARLLDIDETRHADGGKVPIMKMLLESTLRRT 900
 DB 841 LVHRDLAARNVLYKSNHYKITDPFGARLLDIDETRHADGGKVPIMKMLLESTLRRT 900
 QY 901 HOSDVSYGVTWELMTFGAKPYDGIIPAREIPDLKESGERLPOPICTIDVYIMVQCM 960
 DB 901 HOSDVSYGVTWELMTFGAKPYDGIIPAREIPDLKESGERLPOPICTIDVYIMVQCM 960
 QY 961 IDSECRPRRELVSERSMARDPQRFVITONEDLGASPLDSTFFYSLSLEDDMDGLVDA 1020
 DB 961 IDSECRPRRELVSERSMARDPQRFVITONEDLGASPLDSTFFYSLSLEDDMDGLVDA 1020
 QY 1021 EBYLVPQGFPCDPAPAGAGVHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
 DB 1021 EBYLVPQGFPCDPAPAGAGVHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
 QY 1081 AGSDVFDGDLGMAAGLQSLPTHDPSPLOQYSEDETVPLPSETDGYVAPLTCSPQREYV 1140
 DB 1081 AGSDVFDGDLGMAAGLQSLPTHDPSPLOQYSEDETVPLPSETDGYVAPLTCSPQREYV 1140
 QY 1141 NQPDVAPQPPSPREGLPAPAPAGATLERPKTISPGKNGVYKOVFAFGAVENPEYLTPO 1200
 DB 1141 NQPDVAPQPPSPREGLPAPAPAGATLERPKTISPGKNGVYKOVFAFGAVENPEYLTPO 1200
 QY 1201 GGAAPQHPPPAPSPAFNDLYYWDODPPKGAPESTFKGPTAENPEYLGDLVPV 1255
 DB 1201 GGAAPQHPPPAPSPAFNDLYYWDODPPKGAPESTFKGPTAENPEYLGDLVPV 1255

RESULT 2
 ID AAM92406 standard; Protein; 1255 AA.
 XX AAM92406;
 AC AAM92406;
 XX 21-APR-1999 (first entry)
 DT 21-APR-1999
 XX
 DE Human HER-2/neu oncogene protein.
 XX
 KW HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
 KW malignancy; treatment; tumour.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Region
 FT
 PN US5869445-A.
 XX
 XX 09-FEB-1999.
 XX
 PF 01-APR-1996; 96US-0625101.
 XX
 XX 01-APR-1996; 96US-0625101.
 PR 17-MAR-1993; 93US-0033644.
 PR 12-AUG-1993; 93US-0106112.
 PR 31-MAR-1995; 95US-0414417.
 XX
 PA (UNIM) UNIM WASHINGTON.
 XX
 PI Cheever MA, Distie ML;
 XX
 DR MPI; 1999-152835/13.

Location/Qualifiers
 676..1255
 /note="region which elicits immune response"

DR N-PSDB; AAX01912.

XX Use of HER-2/neu polypeptides - for eliciting an immune response to
PT an HER-2/neu associated malignancy, particularly for treating or
PT preventing tumours

PS Claim 3; Column 31-38; 26pp; English.

CC This sequence represents the human HER-2/neu oncogene protein. A fragment
CC of this protein is used in a method for eliciting or enhancing an immune
CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
CC B cells to produce an immune response to the HER-2/neu protein. The
CC method can be used for immunisation against a malignancy in which the
CC HER-2/neu oncogene is associated and in the treatment of an existing
CC tumour, or to prevent tumour occurrence or recurrence.

XX Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 20; Length 1255;

Match Local Similarity 100.0%; Pred. No. 0;

Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELALCRWGLLALPLPGAASVQVCTGDMKRLPASPETHLDMRLHYOGQVYQGNL 60
DB 1 MELALCRWGLLALPLPGAASVQVCTGDMKRLPASPETHLDMRLHYOGQVYQGNL 60
QY 61 ELTYLPTNASLFLDIOEVQGVYLAHNQVQVPLQRLIRVGTQLFEDNTALAVLDNG 120
DB 61 ELTYLPTNASLFLDIOEVQGVYLAHNQVQVPLQRLIRVGTQLFEDNTALAVLDNG 120
QY 121 DPLNNTPTPTGASPGSLRELQRLSLTEILKGVYLRANQQLCYQDTILWKDIFHKNOQLA 180
DB 121 DPLNNTPTPTGASPGSLRELQRLSLTEILKGVYLRANQQLCYQDTILWKDIFHKNOQLA 180
QY 121 DPLNNTPTPTGASPGSLRELQRLSLTEILKGVYLRANQQLCYQDTILWKDIFHKNOQLA 180
DB 121 DPLNNTPTPTGASPGSLRELQRLSLTEILKGVYLRANQQLCYQDTILWKDIFHKNOQLA 180
QY 181 LTLIDTNSRAHCPSPCKSGRCWGSSESDCSLFRYVCAAGSCARCKPGLPTDCHEQC 240
DB 181 LTLIDTNSRAHCPSPCKSGRCWGSSESDCSLFRYVCAAGSCARCKPGLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCYTACP 300
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCYTACP 300
QY 301 YNYLSTDVGSCTLVCPRLHNOVYTABDGTQRCCKSPCARVCYGLAMEHLREYAVTSAN 360
DB 301 YNYLSTDVGSCTLVCPRLHNOVYTABDGTQRCCKSPCARVCYGLAMEHLREYAVTSAN 360
QY 361 IOBPAQCKKIFGSLAFLPESFGDDPASNTAPLOPELOVPELTLEETGLYISAMPDSL 420
DB 361 IOBPAQCKKIFGSLAFLPESFGDDPASNTAPLOPELOVPELTLEETGLYISAMPDSL 420
QY 421 DLSVFNQNLQVIRGRILHNGAYSLTLOGLISWLGRLSLBELSGSLALHNNTHLCVHTV 480
DB 421 DLSVFNQNLQVIRGRILHNGAYSLTLOGLISWLGRLSLBELSGSLALHNNTHLCVHTV 480
QY 481 PWDOLFNNPHQALHTANRPEDECVGEGLAGHQLCARHGWPGPTQVNCOSQPLRGQEC 540
DB 481 PWDOLFNNPHQALHTANRPEDECVGEGLAGHQLCARHGWPGPTQVNCOSQPLRGQEC 540
QY 481 PWDOLFNNPHQALHTANRPEDECVGEGLAGHQLCARHGWPGPTQVNCOSQPLRGQEC 540
DB 481 PWDOLFNNPHQALHTANRPEDECVGEGLAGHQLCARHGWPGPTQVNCOSQPLRGQEC 540
QY 541 VEECRVLOGLPREYVNAHCLFCHPECOFQNGSVTCFGBADQCVAQAHYKDPFVAVAC 600
DB 541 VEECRVLOGLPREYVNAHCLFCHPECOFQNGSVTCFGBADQCVAQAHYKDPFVAVAC 600
QY 601 PSGVPRDLSYMPIMKPPDEBACOPPCINCTHSQVLDLKGCPABORASPLTSISAVVG 660
DB 601 PSGVPRDLSYMPIMKPPDEBACOPPCINCTHSQVLDLKGCPABORASPLTSISAVVG 660
QY 661 ILVVVLGVVFGLILKRAQOKIRKTYMRLLQETELVEPLTSGAMPNPAQRIILKETEL 720
DB 661 ILVVVLGVVFGLILKRAQOKIRKTYMRLLQETELVEPLTSGAMPNPAQRIILKETEL 720
QY 721 RKVKTLGSGARCTYVKGKIWIPIGSENVKILVALKYLRENTSPKANGIILDEAVYMAVGSP 780
DB 721 RKVKTLGSGARCTYVKGKIWIPIGSENVKILVALKYLRENTSPKANGIILDEAVYMAVGSP 780

```

```

QY 781 YVSRLLIGICTSTVQVLTQIMPGCLLDHVENRGLSGQDLNMCQIAKNGSYLEDVR 840
DB 781 YVSRLLIGICTSTVQVLTQIMPGCLLDHVENRGLSGQDLNMCQIAKNGSYLEDVR 840
QY 841 LVHRDLAANNVTVKSNHVKITDFGLARLLDDIETEVHADGGVPIKMALESILRRFT 900
DB 841 LVHRDLAANNVTVKSNHVKITDFGLARLLDDIETEVHADGGVPIKMALESILRRFT 900
QY 901 HSDVMSYVYTWELMTFGAKPYDGI PAREIPDLKEGERLPOPICTIDVYIMVCKM 960
DB 901 HSDVMSYVYTWELMTFGAKPYDGI PAREIPDLKEGERLPOPICTIDVYIMVCKM 960
QY 961 IDSECRPRERELVESESRMARDPORFVLIQNEDLGPASPLDSTFYSLDEDDMGDLVDA 1020
DB 961 IDSECRPRERELVESESRMARDPORFVLIQNEDLGPASPLDSTFYSLDEDDMGDLVDA 1020
QY 1021 ERYLVPQOGFCPPDPAFGAGVHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
DB 1021 ERYLVPQOGFCPPDPAFGAGVHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
QY 1081 AGSDVFDGDLGMAKGLQSLPTHDSPLQRYSEDPVLPSESTDGVAPLTCSPOPEYV 1140
DB 1081 AGSDVFDGDLGMAKGLQSLPTHDSPLQRYSEDPVLPSESTDGVAPLTCSPOPEYV 1140
QY 1141 NQPDVAPQPPSPREBGLPAARPAATLERPKTLSPGKNGVYKDVFAFGAVENPEYLTPO 1200
DB 1141 NQPDVAPQPPSPREBGLPAARPAATLERPKTLSPGKNGVYKDVFAFGAVENPEYLTPO 1200
QY 1201 GGAAPQHPHPAPSPAFDNLTYWDDPPERGAPPESTFGKTPTAENPEYGLDVPV 1255
DB 1201 GGAAPQHPHPAPSPAFDNLTYWDDPPERGAPPESTFGKTPTAENPEYGLDVPV 1255

```

RESULT 3
AAB21198
ID AAB21198 standard; protein; 1255 AA.
AC AAB21198;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human HER-2/neu protein.
XX
KW Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
XX
OS Homo sapiens.
XX
PN MO200044899-A1.
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000MO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.
XX
PA (CORI-) CORIAX CORP.
PA (SMIK-) SMITHLINE BEECHAM.
PI Cheever MA, Gheysen D;
XX
DR MPI: 2000-505976/45.
DR N-PSDB; AAB89736.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Claim 52; Fig 7; 128pp; English.
XX
CC The present sequence is the human HER-2/neu protein. It is a member of

CC the tyrosine kinase family of receptor-like glycoproteins and shows
CC homology to the epidermal growth factor receptor (EGFR). It probably
CC plays a part in cell growth and/or differentiation. The HER-2/neu
CC gene is an oncogene. An HER-2/neu fusion protein comprising a
CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
CC
XX

80 Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 21; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METAAACRMGLLALLPFGAASSTOYCTGTDMKRLPASPETHLDMLHLYOGCQVVOGNTL 60
DB 1 METAAACRMGLLALLPFGAASSTOYCTGTDMKRLPASPETHLDMLHLYOGCQVVOGNTL 60
QY 61 ELTYLPTNASTPFIADIQEYGVYLIANQVROVPLQRLRIYVGTOLFEDNYALVALVDNG 120
DB 61 ELTYLPTNASTPFIADIQEYGVYLIANQVROVPLQRLRIYVGTOLFEDNYALVALVDNG 120
QY 121 DPLNNTPTVPGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMTDIFKNNQLA 180
DB 121 DPLNNTPTVPGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMTDIFKNNQLA 180
QY 121 DPLNNTPTVPGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMTDIFKNNQLA 180
DB 121 DPLNNTPTVPGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMTDIFKNNQLA 180
QY 181 LTIIDTRNSAACHPCSPMCKGSRGWSBSBDCOSLITTVCAAGCARCKGLPTDCHEOC 240
DB 181 LTIIDTRNSAACHPCSPMCKGSRGWSBSBDCOSLITTVCAAGCARCKGLPTDCHEOC 240
QY 181 LTIIDTRNSAACHPCSPMCKGSRGWSBSBDCOSLITTVCAAGCARCKGLPTDCHEOC 240
DB 181 LTIIDTRNSAACHPCSPMCKGSRGWSBSBDCOSLITTVCAAGCARCKGLPTDCHEOC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICELCPALVTNTDTPSPMPNBSGRYTFGASCYACR 300
DB 241 AAGCTGPKHSDCLACHFNHSGICELCPALVTNTDTPSPMPNBSGRYTFGASCYACR 300
QY 301 YNTLSTDVSGSCTIVCPHNOEVTAEQGTORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
DB 301 YNTLSTDVSGSCTIVCPHNOEVTAEQGTORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
QY 361 IOEPFAGCKTFGSLAPFPESFDGDPANTAPLOPBOQVETTLREITGLYISAMPDLP 420
DB 361 IOEPFAGCKTFGSLAPFPESFDGDPANTAPLOPBOQVETTLREITGLYISAMPDLP 420
QY 421 DLSVFOHLOVIRGRILHNGAYSLTLOGISWLGRLSRLSGSLALIHNTHLCFVHTV 480
DB 421 DLSVFOHLOVIRGRILHNGAYSLTLOGISWLGRLSRLSGSLALIHNTHLCFVHTV 480
QY 481 PMDQLFNNPHQALLHTANRPEDCEVSGELACHQUCARHGCWGPPTQCVNCSOPLRGQEC 540
DB 481 PMDQLFNNPHQALLHTANRPEDCEVSGELACHQUCARHGCWGPPTQCVNCSOPLRGQEC 540
QY 541 VESCRVYQGLPREYVNAHCLPCHPBCOPONGSTTCGPRADQCVACHTKDPFCYARC 600
DB 541 VESCRVYQGLPREYVNAHCLPCHPBCOPONGSTTCGPRADQCVACHTKDPFCYARC 600
QY 601 PSQVKPDLSTMPITKPPDEBGACOPCPINCHSCVDLDGKCPAEORASPLTISIISAVG 660
DB 601 PSQVKPDLSTMPITKPPDEBGACOPCPINCHSCVDLDGKCPAEORASPLTISIISAVG 660
QY 661 ILAVVVGAVVGLIKRROQKIRKTYMRRLQETELVEPLTPSGAMPNOQMRLKTEHL 720
DB 661 ILAVVVGAVVGLIKRROQKIRKTYMRRLQETELVEPLTPSGAMPNOQMRLKTEHL 720
QY 721 RKVAVLGSAGFTYKGIWIPDGENVKI PVAIKYLRNENTSFKANKEILDEAYVAGVGP 780
DB 721 RKVAVLGSAGFTYKGIWIPDGENVKI PVAIKYLRNENTSFKANKEILDEAYVAGVGP 780
QY 781 YVSLHLDGILSTVQVLTQMLPVGCLLDHYENNGRLSGODLLMWCMQIAKMSYLDVR 840
DB 781 YVSLHLDGILSTVQVLTQMLPVGCLLDHYENNGRLSGODLLMWCMQIAKMSYLDVR 840

QY 841 LVHRDIAANVLYKSNHYKTFDGLARLLIDETSYHADGKVPKIMMALESILRRRFT 900
DB 841 LVHRDIAANVLYKSNHYKTFDGLARLLIDETSYHADGKVPKIMMALESILRRRFT 900
QY 901 HOSDVSXVYTWELMTFPAKPYDGI PAREIPDLKEGRLPDPICITDVYIMVCM 960
DB 901 HOSDVSXVYTWELMTFPAKPYDGI PAREIPDLKEGRLPDPICITDVYIMVCM 960
QY 961 ISECRPRELVSERSMARDPQRFVIONEDLGPASPLDSTFYRSLYEDMDGLVDA 1020
DB 961 ISECRPRELVSERSMARDPQRFVIONEDLGPASPLDSTFYRSLYEDMDGLVDA 1020
QY 1021 ERTLVQGFPCDPAPGAGVYHHRSSSTRSGGDLTLGLEPSEEARPSPLASEG 1080
DB 1021 ERTLVQGFPCDPAPGAGVYHHRSSSTRSGGDLTLGLEPSEEARPSPLASEG 1080
QY 1081 AGSDVFDGDLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVAPLTCSPQREYV 1140
DB 1081 AGSDVFDGDLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVAPLTCSPQREYV 1140
QY 1141 NQPDVAPQPPSPREGPLPAPRAGATLERPKTISPKNQGVKDVAFGAVENPEYLTPQ 1200
DB 1141 NQPDVAPQPPSPREGPLPAPRAGATLERPKTISPKNQGVKDVAFGAVENPEYLTPQ 1200
QY 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPPERGAPSTFKGTPTABNPEYLGLDVPV 1255
DB 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPPERGAPSTFKGTPTABNPEYLGLDVPV 1255

RESULT 4
AAV84780
ID AAV84780 standard; Protein; 1255 AA.
XX
AC AAV84780;
XX
DT 08-ANG-2000 (first entry)
XX
DE Amino acid sequence of the SPLICE erbB-2 receptor protein.

XX
XX SPLICE erbB-2 receptor protein; cell transformation disorder; cancer;
KW tumor cell proliferation; tissue degeneration; atrophy;
KW bone resorption; inflammatory disease; degenerative disorder;
KW wound healing.
OS
XX Homo sapiens.
XX
PN M0200020579-A1.
XX
PD 13-APR-2000.
XX
PF 01-OCT-1999; 99MO-CA00912.
XX
PR 02-OCT-1998; 98US-0165192.
XX
PA (UIMC-) UNIV MCMASTER.
XX
PI Muller WJ, Siegel PM;
XX
DR WPI, 2000-103768/26.
XX
DR N-PSDB; AAA14812.
XX

PT Nucleic acid encoding an erbB-2 receptor protein designated SPLICE
PT erbB-2, inhibitors of the protein are useful for treatment of cancer -
XX
XX
PS Claim 3; Fig 2; 60pp; English.

CC The present sequence represents a SPLICE erbB-2 receptor protein. The
CC protein has an in-frame deletion of 16 amino acids, 2 of which are
CC conserved cysteine residues, compared to the unspliced protein. The
CC erbB-2 polynucleotide is used to construct probes for detecting
CC disorders of cell transformation such as cancer. Antibodies to the
CC protein may be used to detect SPLICE erbB-2 in a sample. Agents
CC (e.g. antisense oligonucleotides) which inhibit the expression of

CC Splice erbb-2 are useful for reducing tumor cell proliferation and
 CC treating cancer. Substances which stimulate Splice erbb-2 are useful
 CC for treating conditions involving damaged cells including conditions
 CC in which degeneration of tissue occurs, such as arthropathy, bone
 CC resorption, inflammatory diseases, degenerative disorders of the
 CC central nervous system and wound healing.

XX Sequence 1255 AA:

Query Match 100.0%; Score 6815; DB 21; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPPGAASVCTGTDMKALPASPEETHLMDRLHLYGCGVQGNL 60
 DB 1 MELAALCRWGLLALLPPGAASVCTGTDMKALPASPEETHLMDRLHLYGCGVQGNL 60
 QY 61 ELTYLPTNASLSPLODIQEVGVYLIAHNOVROVPLQRLRTVRGTOLPEDNYALAVDNG 120
 DB 61 ELTYLPTNASLSPLODIQEVGVYLIAHNOVROVPLQRLRTVRGTOLPEDNYALAVDNG 120
 QY 121 DPLNNTPTVGAASPGRLRLQRLSLTEILKGVLIORNPOLCYODTILMKDIIFKNNOLA 180
 DB 121 DPLNNTPTVGAASPGRLRLQRLSLTEILKGVLIORNPOLCYODTILMKDIIFKNNOLA 180
 QY 181 LTLIDTNRSPACHSCFPMCKGSRGSSSDQSLTRTVAGGACRCKPLPTDCHEQC 240
 DB 181 LTLIDTNRSPACHSCFPMCKGSRGSSSDQSLTRTVAGGACRCKPLPTDCHEQC 240
 QY 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTNTDTPESMNPBEGRYTFGASCTTAC 300
 DB 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTNTDTPESMNPBEGRYTFGASCTTAC 300
 QY 301 YNTLSTVSGCTVACPLHNOEVTAEDEGTORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
 DB 301 YNTLSTVSGCTVACPLHNOEVTAEDEGTORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
 QY 361 IQEPAGCKKI FGSIALFLPESFDGDPASNTAPLOEQLQVETLEITGYLYISAMPDLP 420
 DB 361 IQEPAGCKKI FGSIALFLPESFDGDPASNTAPLOEQLQVETLEITGYLYISAMPDLP 420
 QY 421 DLSVPQNLQYIRGRILHNGAYSLTLOGLSLGLRLRLGSGALAIHNTHLCPYHTV 480
 DB 421 DLSVPQNLQYIRGRILHNGAYSLTLOGLSLGLRLRLGSGALAIHNTHLCPYHTV 480
 QY 481 PMDOLFNPHQALLHTANRPEDECVGEGIALCHQICARHGMCPPTQCVNCSQPLRQEC 540
 DB 481 PMDOLFNPHQALLHTANRPEDECVGEGIALCHQICARHGMCPPTQCVNCSQPLRQEC 540
 QY 541 VEECRVLOGILPREYVNAHCLPCHPECOQONGSVTCFGEPAQOCVACAHYKDPFCVARC 600
 DB 541 VEECRVLOGILPREYVNAHCLPCHPECOQONGSVTCFGEPAQOCVACAHYKDPFCVARC 600
 QY 601 PEGVAKDLSMPYWKPFDEGACQPCPINCTHSCVDLIDDKCPAEGORASPLTISIAYVG 660
 DB 601 PEGVAKDLSMPYWKPFDEGACQPCPINCTHSCVDLIDDKCPAEGORASPLTISIAYVG 660
 QY 661 ILLVVLGVVFGILIKRROOKIRKYMTRRLQETELVBPILPSGAMNOQMLKTEL 720
 DB 661 ILLVVLGVVFGILIKRROOKIRKYMTRRLQETELVBPILPSGAMNOQMLKTEL 720
 QY 721 RKVKVLSGAGFYVYKGIWIPDGENVKI PVAILKVLRENTSPKANKELIDDAVYVAGGSP 780
 DB 721 RKVKVLSGAGFYVYKGIWIPDGENVKI PVAILKVLRENTSPKANKELIDDAVYVAGGSP 780
 QY 781 YVSRLLGICLTSTVQVLTQMLPGCLLDHYRENRKRLGSDLLNMCQIAKGSYLEDVR 840
 DB 781 YVSRLLGICLTSTVQVLTQMLPGCLLDHYRENRKRLGSDLLNMCQIAKGSYLEDVR 840
 QY 841 LVHRDLAABNVLYKSPHVKITDGRALRLIDIDETEHADGKGVPIKMALESTLRRT 900
 DB 841 LVHRDLAABNVLYKSPHVKITDGRALRLIDIDETEHADGKGVPIKMALESTLRRT 900

QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIAPAREIPDLLEKGERLPQPPICITIDVYMINVKCM 960
 DB 901 HQSDVMSYGVTVWELMTFGAKPYDGIAPAREIPDLLEKGERLPQPPICITIDVYMINVKCM 960
 QY 961 IDSECRPRFRELVEFSRMARDPQRFVYIQWEDLGPASPLDSTFPRSLIEDDMGDLYDA 1020
 DB 961 IDSECRPRFRELVEFSRMARDPQRFVYIQWEDLGPASPLDSTFPRSLIEDDMGDLYDA 1020
 QY 1021 EBYLVPOQGFPCPPAGAGGMVHRHRSSTRSGGDLTGLFSESEAPRSLAPSEB 1080
 DB 1021 EBYLVPOQGFPCPPAGAGGMVHRHRSSTRSGGDLTGLFSESEAPRSLAPSEB 1080
 QY 1081 AGSDVPDGLGMAKGLQSLPTHDPSPLORYSDPTVLPSETDGYVAPLTCSQPEVY 1140
 DB 1081 AGSDVPDGLGMAKGLQSLPTHDPSPLORYSDPTVLPSETDGYVAPLTCSQPEVY 1140
 QY 1141 NQPDVRQPPSPRSGPLPAARPAATLERPKTSLPGKGVVXVYPAFGAVENPEYLTPO 1200
 DB 1141 NQPDVRQPPSPRSGPLPAARPAATLERPKTSLPGKGVVXVYPAFGAVENPEYLTPO 1200
 QY 1201 GGAAPQPPPPAPSPAFDNLVYMDQDPPBRGAPSTFGPTANPEYLGIDVPV 1255
 DB 1201 GGAAPQPPPPAPSPAFDNLVYMDQDPPBRGAPSTFGPTANPEYLGIDVPV 1255

RESULT 5
 ID AAB85458 standard; protein; 1255 AA.
 XX AAB85458;
 AC AAB85458;
 DT 25-SEP-2001 (first entry)
 XX
 DE Human HER-2/neu protein.
 XX
 KW Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
 KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2.
 XX
 OS Homo sapiens.
 XX
 PN W0200153463-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 19-JAN-2001; 2001MO-US01850.
 XX
 PR 21-JAN-2000; 2000US-0177545.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Cheever MA, Hand-Zimmermann S;
 XX
 DR WPI, 2001-476112/51.
 DR N-PSDB; AAH23392.
 XX
 PT New antigen-presenting cells, useful as vaccines for eliciting or
 PT enhancing an immune response to HER-2/neu protein, particularly useful
 PT for treating or preventing cancer, e.g. breast cancer
 XX
 PS Claim 2; Page 41-46; 49p; English.
 XX
 CC The invention provides an isolated antigen-presenting cell, which
 CC expresses at least an immunogenic portion of a polypeptide that produces
 CC an immune response to HER-2/neu protein. The antigen-presenting cells are
 CC useful as vaccines for eliciting or enhancing an immune response to
 CC HER-2/neu protein, particularly in treating or preventing malignancies in
 CC which the HER-2/neu oncogene is associated. Specifically, these are
 CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
 CC colon, lung or prostate cancers. The present sequence represents
 CC the human HER-2/neu protein (also known as p185 or c-erbB2).
 XX
 SQ Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 22; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 METALALCRMGGLLALLLPAGAASTOVCTGTBMKRLPASPTHLDMRLHYOGCCOVGNL 60
 1 METALALCRMGGLLALLLPAGAASTOVCTGTBMKRLPASPTHLDMRLHYOGCCOVGNL 60
 61 ELTYLPTNASLSFLQDIQEVGYVLIANOVQVPLRLRVSTQVLFEDNYALAVLDNG 120
 61 ELTYLPTNASLSFLQDIQEVGYVLIANOVQVPLRLRVSTQVLFEDNYALAVLDNG 120
 121 DPLANTPTVYVGAAGGRLRLQRLSLTEILKGVLIQGNPOLCYDITLMDIPIKNNOLA 180
 121 DPLANTPTVYVGAAGGRLRLQRLSLTEILKGVLIQGNPOLCYDITLMDIPIKNNOLA 180
 181 LTLIDTRSRACCHSCPMCKGSRCKMGSSSEDCQSLTITVCAGGCAKCKGFLPTDCCHEOC 240
 181 LTLIDTRSRACCHSCPMCKGSRCKMGSSSEDCQSLTITVCAGGCAKCKGFLPTDCCHEOC 240
 241 AAGCTGPKHSDCLAELPHNSGICELHCPALVTYNTDTPESMPNBSGYTFGASCVTACP 300
 241 AAGCTGPKHSDCLAELPHNSGICELHCPALVTYNTDTPESMPNBSGYTFGASCVTACP 300
 301 YNTLSTVGSCTLVCPILHNOETVRAEDGTORCEKSKFCARVCTGLGMEHLREVRATVSAN 360
 301 YNTLSTVGSCTLVCPILHNOETVRAEDGTORCEKSKFCARVCTGLGMEHLREVRATVSAN 360
 361 IOEPAGCKKIFGSLAPLPESFDPGPAANTAPLOEPOLOVPTLEBITGYLISAMPISLP 420
 361 IOEPAGCKKIFGSLAPLPESFDPGPAANTAPLOEPOLOVPTLEBITGYLISAMPISLP 420
 421 DLSVPONLQVIRGRILHNGAYSLTQGLI9WLGRLRLRELSGSLALIHNTHLCPVHTV 480
 421 DLSVPONLQVIRGRILHNGAYSLTQGLI9WLGRLRLRELSGSLALIHNTHLCPVHTV 480
 481 PMOOLPNNPHOALHTANRPEDCEVGBGLACHOUCAHGHCMPPTQCNCSOPLRODEC 540
 481 PMOOLPNNPHOALHTANRPEDCEVGBGLACHOUCAHGHCMPPTQCNCSOPLRODEC 540
 541 VEECRVLOGLPREVYNNARHCLPCHPECOPONGSYTCFGEPAADQVACAHYKDPFCVARC 600
 541 VEECRVLOGLPREVYNNARHCLPCHPECOPONGSYTCFGEPAADQVACAHYKDPFCVARC 600
 601 PSQVXPDLSYMP1WKFPDEBGACQPCPINCTHSCVDLDKGCAPAGORASPLTISAVG 660
 601 PSQVXPDLSYMP1WKFPDEBGACQPCPINCTHSCVDLDKGCAPAGORASPLTISAVG 660
 661 ILVVVLTGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAMRILKETEEL 720
 661 ILVVVLTGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAMRILKETEEL 720
 721 RKVYVLSGAGFTYVKGIMIPDGENVKIPVAIKYLRNTSPKAKKILDEAYVWAGVSP 780
 721 RKVYVLSGAGFTYVKGIMIPDGENVKIPVAIKYLRNTSPKAKKILDEAYVWAGVSP 780
 781 YVSRLLGICLTSTVOLVTOIMPYCLLDHVENRGRLSQDLMWCMQIAGMSYLEDVR 840
 781 YVSRLLGICLTSTVOLVTOIMPYCLLDHVENRGRLSQDLMWCMQIAGMSYLEDVR 840
 841 LVHRDLAARVNVLSVSPHVKITDGLARLLDIDETEVHADGKVP1KMMALSLIRRRFT 900
 841 LVHRDLAARVNVLSVSPHVKITDGLARLLDIDETEVHADGKVP1KMMALSLIRRRFT 900
 901 HOSDVMISGYVTWMLMTFGAKPYDIPARBEIPDLBKGEBRLPOPPICITIDVYMMVCM 960
 901 HOSDVMISGYVTWMLMTFGAKPYDIPARBEIPDLBKGEBRLPOPPICITIDVYMMVCM 960
 961 IDSECRPRFRELVESEFARMADPORFVVIQNEBGLGPA5PLDSTFYRSLLEDMDGDLVDA 1020
 961 IDSECRPRFRELVESEFARMADPORFVVIQNEBGLGPA5PLDSTFYRSLLEDMDGDLVDA 1020
 1021 BEYLVPOQGFPCDDPAPAGAGVHHRRSSSTRSGGGLTLGLPSSSEBAFRSLAPSE 1080

1021 BEYLVPOQGFPCDDPAPAGAGVHHRRSSSTRSGGGLTLGLPSSSEBAFRSLAPSE 1080
 1081 AGSDVDFDGLGMAAGKLOSLPTHDPSPLOXYSEDPVPLPSESTDGVAAFLTCSPOBEYV 1140
 1081 AGSDVDFDGLGMAAGKLOSLPTHDPSPLOXYSEDPVPLPSESTDGVAAFLTCSPOBEYV 1140
 1141 NQPDVRRPQPSPEEGULPAARPAATLERPKTISPGKNGYVAKVFAFGAVENPEYLTPO 1200
 1141 NQPDVRRPQPSPEEGULPAARPAATLERPKTISPGKNGYVAKVFAFGAVENPEYLTPO 1200
 1201 GGAAPQHPPEPASPAPFDMLYYWDODPERRGAPSTFKGPTAENPEYLGIDVAV 1255
 1201 GGAAPQHPPEPASPAPFDMLYYWDODPERRGAPSTFKGPTAENPEYLGIDVAV 1255

RESULT 6
 ID AAG88267 standard; Protein; 1255 AA.
 XX AAG88267;
 AC AAG88267;
 XX AAG88267;
 DT 11-SEP-2001 (first entry)
 XX 11-SEP-2001 (first entry)
 DE HER2/neu amino acid sequence.
 XX Human; HER2/neu; epitope; human leukocyte antigen; HLA; T cell;
 KM immune response; vaccine; tumour; cancer; cytotoxic; immunostimulant;
 KM tumour-associated antigen; T lymphocyte; cytotoxic T lymphocyte; CTL.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200141787-A1.
 XX 14-JUN-2001.
 PD 14-JUN-2001.
 PF 11-DEC-2000; 2000MO-US33591.
 XX 11-DEC-2000; 2000MO-US33591.
 XX 10-DEC-1999; 99US-0458299.
 PR 10-DEC-1999; 99US-0458299.
 XX (EPIM-) EPIMMUNE INC.
 PA (EPIM-) EPIMMUNE INC.
 PI Pike J, Sette A, Sidney J, Southwood S, Chesnut R, Celis E;
 PI Keogh E;
 DR WPI; 2001-374995/39.
 XX WPI; 2001-374995/39.
 PT An isolated prepared HER2/neu epitope useful in a vaccine for inducing
 PT cellular immune responses for the prevention and treatment of cancer -
 PS Disclosure; Page 15; 1999; English.
 XX The present invention describes isolated prepared HER2/neu epitopes (I).
 XX Also described are: (1) a clonal cytotoxic T lymphocyte (CTL) that is
 XX culture in vitro and binds to a complex of an epitope (I), bound to a
 XX human leukocyte antigen (HLA) molecule; (2) a peptide (II) comprising (I)
 XX and a second epitope and the peptide is less than 50 contiguous amino
 XX acids that have 100% identity with a native peptide sequence of HER2/neu;
 XX (3) a vaccine composition (III) comprising (II) and a pharmaceutical
 XX excipient; (4) an isolated nucleic acid encoding a peptide comprising
 XX (I); and (5) an isolated nucleic acid encoding (II). (I) has cytotoxic
 XX and immunostimulant activities, and can be used in vaccines. (I), (II)
 XX and (III) are useful for inducing cellular immune responses for the
 XX prevention and treatment of cancer. (I) and (II) are useful for
 XX monitoring or evaluating an immune response to a tumour-associated
 XX antigen when incubated with a T lymphocyte sample from a patient and
 XX detecting the presence of bound T lymphocyte to (I) or (II). Epitope
 XX based vaccines mean that immunosuppressive epitopes that may be present
 XX in whole antigens may be avoided. Selected epitopes may be combined to
 XX enhance immunogenicity. The possible pathological side effects caused by
 XX infectious agents or whole protein antigen is eliminated. The vaccine
 XX provides the ability to direct and focus an immune response to multiple
 XX selected antigens from the same pathogen. Epitope-based anti-tumour

CC vaccines provides the opportunity to combine epitopes derived from
 CC multiple tumour-associated molecules addressing the problem of tumour-
 CC tumour variability and reducing the likelihood of tumour escape due to
 CC antigen loss. AAG88266 to AAG89121 represent amino acid sequences used in
 CC the exemplification of the present invention.

Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 22; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 METALACRMGLALALPRGAASQVCTGDMKRLPASPEHMLDMLEHYOGCQVQGNL 60
 1 METALACRMGLALALPRGAASQVCTGDMKRLPASPEHMLDMLEHYOGCQVQGNL 60
 61 ELTYLPTNASLSFLQDIQEVGYVLIANQVRQVPLRLIVRGTLFEDNYALAVDNG 120
 61 ELTYLPTNASLSFLQDIQEVGYVLIANQVRQVPLRLIVRGTLFEDNYALAVDNG 120
 121 DPANNTTPTVAGAPGGRLRLQLRLTEILKGVLIQGNPOLCYQDTILMDIFHKNQLA 180
 121 DPANNTTPTVAGAPGGRLRLQLRLTEILKGVLIQGNPOLCYQDTILMDIFHKNQLA 180
 181 LTLIDNRSBACHPCSPMCKSGRCWSSSDCSLTFTVCAAGCARCKGRLPTDCCHEOC 240
 181 LTLIDNRSBACHPCSPMCKSGRCWSSSDCSLTFTVCAAGCARCKGRLPTDCCHEOC 240
 241 AAGCTPKISDCLACIHFHNSGICELHCPALVTYNTDTPSMNPBGRYTFGASCYACP 300
 241 AAGCTPKISDCLACIHFHNSGICELHCPALVTYNTDTPSMNPBGRYTFGASCYACP 300
 301 YNYLSTDVSGCTLVCPRLNQEVTAEDGTORCEKCSKPCARVCTGLGMEHLREYRAVTSAN 360
 301 YNYLSTDVSGCTLVCPRLNQEVTAEDGTORCEKCSKPCARVCTGLGMEHLREYRAVTSAN 360
 361 IOEPAGCKKIFGSLALPSPFPDDPASNTAPLPOLQVETLEITGYLYISAMPDSL 420
 361 IOEPAGCKKIFGSLALPSPFPDDPASNTAPLPOLQVETLEITGYLYISAMPDSL 420
 421 DLSVFONLQVIRGRILHANGAYSLTLQGLISWLGSLRLSLRLSGALLIHNTHLCEVHTV 480
 421 DLSVFONLQVIRGRILHANGAYSLTLQGLISWLGSLRLSLRLSGALLIHNTHLCEVHTV 480
 481 PWDOLFERNPHOALLHTANRPEDECVGEGLACHQLCARGHCWGPPTQCVNCSQFLRGSEC 540
 481 PWDOLFERNPHOALLHTANRPEDECVGEGLACHQLCARGHCWGPPTQCVNCSQFLRGSEC 540
 541 VEECRVLOGLPRBYNARHCLPCHPECCQPNQSVTCFGBPADQCVACAHYKDPFCVABC 600
 541 VEECRVLOGLPRBYNARHCLPCHPECCQPNQSVTCFGBPADQCVACAHYKDPFCVABC 600
 601 PSGVNPDLSTYMPILWKPDEEGACQPCINCTHSCVLDKGCAGABRASPLTISIAYVG 660
 601 PSGVNPDLSTYMPILWKPDEEGACQPCINCTHSCVLDKGCAGABRASPLTISIAYVG 660
 661 ILVVVVLGVVFGSILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAMRLKETEL 720
 661 ILVVVVLGVVFGSILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAMRLKETEL 720
 721 RKVKVLSGAGFTVYKGIWIPDGENVKIPVAIKVLAENTSPKANKELIDEAAYMAGVGP 780
 721 RKVKVLSGAGFTVYKGIWIPDGENVKIPVAIKVLAENTSPKANKELIDEAAYMAGVGP 780
 781 YVSRLLGICLTSTVQLVQLMPYGLLDVYREKRGSLGSDLIANWCQIAKMSYLEDVR 840
 781 YVSRLLGICLTSTVQLVQLMPYGLLDVYREKRGSLGSDLIANWCQIAKMSYLEDVR 840
 841 LVHRDLAARNVAVKSPNHAKITDFGLARLLDIDETEHADGGKVPKMMALSTIILRRPT 900
 841 LVHRDLAARNVAVKSPNHAKITDFGLARLLDIDETEHADGGKVPKMMALSTIILRRPT 900
 901 HOSDVMSYGVTVWELMTFGAKPYDGI PARBIDPLEKGERLPOPPICTIDVYMIWKCMM 960

DB 901 HOSDVMSYGVTVWELMTFGAKPYDGI PARBIDPLEKGERLPOPPICTIDVYMIWKCMM 960
 QY 961 IDSECRFRELVESEFRMARDPQRFVIONEDIGPASPLDSTYRSLLEDDMDGLVDA 1020
 DB 961 IDSECRFRELVESEFRMARDPQRFVIONEDIGPASPLDSTYRSLLEDDMDGLVDA 1020
 QY 1021 EBYLVPOQGFPCDPAPAGAGVYHRRSSSTSGGADLTGLPSESEARSPPLAPSEG 1080
 DB 1021 EBYLVPOQGFPCDPAPAGAGVYHRRSSSTSGGADLTGLPSESEARSPPLAPSEG 1080
 QY 1081 AGSDVDFDGLGKAAKGLSLPTHDSPLQRYSEDPVPLPSETDGYVAPLTCSPOPEYV 1140
 DB 1081 AGSDVDFDGLGKAAKGLSLPTHDSPLQRYSEDPVPLPSETDGYVAPLTCSPOPEYV 1140
 QY 1141 NQPDVAPQPPSPREGPLPARPAGATLERPKTSLPGKNGVVKVYFAGAVENPEYLTPO 1200
 DB 1141 NQPDVAPQPPSPREGPLPARPAGATLERPKTSLPGKNGVVKVYFAGAVENPEYLTPO 1200
 QY 1201 GGAAPQHPPEPARSPAFDNLVYWDOPPERGAPSPTEKGTPTAENPEYLGIDVPV 1255
 DB 1201 GGAAPQHPPEPARSPAFDNLVYWDOPPERGAPSPTEKGTPTAENPEYLGIDVPV 1255

RESULT 7

AAE24067 standard; Protein; 1255 AA.

AAE24067;
 23-SEP-2002 (first entry)
 Human Her-2 protein.
 Human; Her-2; epidermal growth factor receptor 2; infection; cancer;
 hyperproliferative disorder; prophylaxis; inflammation; antisense;
 tumour; gene therapy; phosphorothioate backbone.
 Homo sapiens.
 WO200222636-A1.
 21-MAR-2002.
 12-SEP-2001; 2001WO-US28572.
 15-SEP-2000; 2000US-0663834.
 (ISIS-) ISIS PHARM INC.
 Bennett CF, Cowseert LM;
 WPI; 2002-471192/50.
 DR N-PSDB; AAD88904.
 Novel antisense oligonucleotide which modulates the expression of Human
 Epidermal Growth Factor receptor, Her2, is useful for treating tumors
 inflammation or to prevent infection in humans -
 Example 13; Page 95-107; 116pp; English.
 The invention relates to antisense compounds targeted to a nucleic
 acid molecule encoding Her2 (human Epidermal Growth Factor receptor 2)
 that specifically hybridises with and inhibits the expression of Her2.
 Antisense compounds of the invention are used for treating diseases or
 conditions associated with Her2 such as hyperproliferative disorders
 e.g. lung, breast, gastric, oesophageal, colon, bladder, salivary,
 neural or cardiac cancer. They are also useful prophylactically e.g.
 to prevent or delay infection, inflammation and tumour formation. The
 invention is also used in gene therapy. The present sequence is human
 Her-2 protein.
 Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 23; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MELAALCRWGLLALPLPAGASTVCTGTMTKRLPASPTHMLMLHLYOGCOVGNL 60
 1 MELAALCRWGLLALPLPAGASTVCTGTMTKRLPASPTHMLMLHLYOGCOVGNL 60
 61 ELTYLPTNALSIFLODIOEVGVYLIANOVQVPLRLRIVRTQLEFENYALAVDNG 120
 61 ELTYLPTNALSIFLODIOEVGVYLIANOVQVPLRLRIVRTQLEFENYALAVDNG 120
 121 DPLANTTPTVGSAGGRLRELOLSSTILKGVLIQNNPOLCYDOTTLMKDIFFKNNOLA 180
 121 DPLANTTPTVGSAGGRLRELOLSSTILKGVLIQNNPOLCYDOTTLMKDIFFKNNOLA 180
 181 LTLIDNRSRACHPCSPCKGSRCKWSSSDCOSLTRTVACGACGCKPLPDDCCEOC 240
 181 LTLIDNRSRACHPCSPCKGSRCKWSSSDCOSLTRTVACGACGCKPLPDDCCEOC 240
 241 AACCTGPKSDCLACLFHNSGICELHCPALVTYNTDTPESMPNBSRYTFGASCTVACP 300
 241 AACCTGPKSDCLACLFHNSGICELHCPALVTYNTDTPESMPNBSRYTFGASCTVACP 300
 301 YNLTSTDVSGCTIVCPILHNOEVTAEQTCCKSKPCANVCYGLGMEHLREVAVTSAN 360
 301 YNLTSTDVSGCTIVCPILHNOEVTAEQTCCKSKPCANVCYGLGMEHLREVAVTSAN 360
 361 IOEPAGCKKI FGSIAFLPESPDGPASNTAPLOEBOLOVEETLEITGYLISMPDLP 420
 361 IOEPAGCKKI FGSIAFLPESPDGPASNTAPLOEBOLOVEETLEITGYLISMPDLP 420
 421 DLSVFNQLQVIRGRILHNGAYSLTLOGLISWLGRLREISGIALIHNHNLCPHNTV 480
 421 DLSVFNQLQVIRGRILHNGAYSLTLOGLISWLGRLREISGIALIHNHNLCPHNTV 480
 481 PMDQLFNPQHIALHTNRPBDECVGEGLAACHOLCARGCKGPGPTCCVCSOPLRGQEC 540
 481 PMDQLFNPQHIALHTNRPBDECVGEGLAACHOLCARGCKGPGPTCCVCSOPLRGQEC 540
 541 VESCRVLOGLPREVYNARHCLPCHBEGOPONGSVTCGPRADQCAAHKDPFCVARG 600
 541 VESCRVLOGLPREVYNARHCLPCHBEGOPONGSVTCGPRADQCAAHKDPFCVARG 600
 601 PSGVKEPDLSTYMPILKFPDEBEGACQPCPINCSTHCVDDLDKCPABQASPLTSLISAVG 660
 601 PSGVKEPDLSTYMPILKFPDEBEGACQPCPINCSTHCVDDLDKCPABQASPLTSLISAVG 660
 661 ILVVVVGAVFGILIKRQOKIRKRYNRRLIQETELVEPLTPSGAMPNOAMRLKETEL 720
 661 ILVVVVGAVFGILIKRQOKIRKRYNRRLIQETELVEPLTPSGAMPNOAMRLKETEL 720
 721 RKRYVGGSGAGVYVKGIMT PDGENVKIPVALIKYLRNTSPKAKELIDRAYVNAAGVSP 780
 721 RKRYVGGSGAGVYVKGIMT PDGENVKIPVALIKYLRNTSPKAKELIDRAYVNAAGVSP 780
 781 YVSRLLGICLTSTVQVLTQMLPQCLDHYENRGRIGSODLNMCMQIAKMSYLEDVR 840
 781 YVSRLLGICLTSTVQVLTQMLPQCLDHYENRGRIGSODLNMCMQIAKMSYLEDVR 840
 841 LVHRDLAARVNLVKSPPHVKITDFGLARLLDIDETBYHADGKVPKIMMALBETLRRFT 900
 841 LVHRDLAARVNLVKSPPHVKITDFGLARLLDIDETBYHADGKVPKIMMALBETLRRFT 900
 901 HOSDWSYGVTVWELMTFGAKPYDGIPIAREIPDLLEBGERLPOPPICITIVYMMVCM 960
 901 HOSDWSYGVTVWELMTFGAKPYDGIPIAREIPDLLEBGERLPOPPICITIVYMMVCM 960
 961 IDSECRFRFELVSEFRMARDPQPVYIQNEDLGPASPLDSTYRSLLDEDMDGLVDA 1020
 961 IDSECRFRFELVSEFRMARDPQPVYIQNEDLGPASPLDSTYRSLLDEDMDGLVDA 1020

QY 1021 EELTVPOQGFPCDPAPAGAGVHHRRSSSTSSGGDLTLGLRPSHEAPRSLAPSEG 1080
 DB 1021 EELTVPOQGFPCDPAPAGAGVHHRRSSSTSSGGDLTLGLRPSHEAPRSLAPSEG 1080
 QY 1081 AGSDVFDGDLGMAAGLOSLPTHDPSPLOQYSEDPVPLPSETDGVVAPLTCSPOPEYV 1140
 DB 1081 AGSDVFDGDLGMAAGLOSLPTHDPSPLOQYSEDPVPLPSETDGVVAPLTCSPOPEYV 1140
 QY 1141 NQPDVBPQPSPREGLPAPAPAGATLERPKTSLPGNGVVKOVFAFGAVENBEYLTPQ 1200
 DB 1141 NQPDVBPQPSPREGLPAPAPAGATLERPKTSLPGNGVVKOVFAFGAVENBEYLTPQ 1200
 QY 1201 GGAAPQHPPEPASPAPNDLYYNDOPPERGAPSPSTFKGPTAENPEYLGIDVAV 1255
 DB 1201 GGAAPQHPPEPASPAPNDLYYNDOPPERGAPSPSTFKGPTAENPEYLGIDVAV 1255

RESULT 8
 AAE20479
 ID AAE20479 standard; Protein; 1255 AA.
 XX
 AC AAE20479;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Human Her-2/neu protein.
 XX
 KM Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 KM human leukocyte antigen; HLA; vaccine; malignancy; cytostatic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1021..1030
 XX /note="Naturally processed HLA-B44-restricted epitope"
 XX
 PN NO200214503-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 14-AUG-2001; 2001MO-US41733.
 XX
 PR 14-AUG-2000; 2000US-225152P.
 PR 28-SEP-2000; 2000US-236428P.
 PR 21-FEB-2001; 2001US-270520P.
 XX
 PA (CORI-) CORIAX CORP.
 XX
 PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
 PI McNeill PD, Vedvick TS;
 DR WPI; 2002-280758/32.
 DR N-PSDB; AAD32743.
 XX
 PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 PT prevention and diagnosis of cancer, preferably breast cancer -
 XX
 PS Disclosure; Page 114-117; 129pp; English.
 XX
 CC The invention relates to an isolated Her-2/Neu polypeptide composition
 CC effective for eliciting an immune response. The invention is useful for
 CC eliciting an immune response in a patient, where the patient is human
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
 CC The composition is useful for the therapy and diagnosis of cancer,
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
 CC and other compositions for the diagnosis, prevention and treatment of
 CC human malignancies, for stimulating and/or expanding T cells specific for
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 CC patient. The invention is useful for stimulating a T cell response in a
 CC human patient, as probe or primer for nucleic acid hybridisation, to
 CC selectively form duplex molecules with complementary stretches of the
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
 CC length gene from a suitable library, and to direct expression of a

in a variety of cancers, including breast, ovarian, colon, lung and prostate cancer. Her-2/neu is a member of the tyrosine kinase family of receptor-like glycoproteins. It comprises an extracellular domain with homology to the epidermal growth factor receptor (EGFR), a highly hydrophobic transmembrane domain and a C-terminal intracellular domain that also shows homology to EGFR. Its overexpression correlates with a poor prognosis in breast and ovarian cancers. The invention provides Her-2/neu fusion proteins, nucleic acids encoding them, viral vectors, and vaccines comprising the fusion proteins or nucleic acid molecules. In preferred fusion proteins, the extracellular domain of a Her-2/neu protein is fused to a Her-2/neu intracellular domain or a phosphorylation domain (or its DelatpD fragment). An immune response to Her-2/neu protein is elicited or enhanced by administering the fusion protein in the form of a vaccine, or by transfecting cells of an animal *ex vivo* with a nucleic acid encoding the fusion protein, and delivering the transfected cells to the animal. The fusion proteins, nucleic acids, and isolated specific T-cells are useful for inhibiting the development of a cancer, especially breast, ovarian, colon, lung or prostate cancer in a patient. T cells that specifically react with a Her-2/neu fusion protein can be used to remove tumour cells from a sample in order to inhibit the development of cancer in a patient.

Sequence 1255 AA;
Query Match 100.0%; Score 6815; DB 23; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MELALALCRWGLLALLPRGAASVQVCTGTMKLLPASPETHLMLEHLVGCQVQGNL 60
1 MELALALCRWGLLALLPRGAASVQVCTGTMKLLPASPETHLMLEHLVGCQVQGNL 60
61 ELTLPTNASTLPQDIOEVGVYLIANQVQVPLRLRIVRSTQPLFEDNYALAVLDNG 120
61 ELTLPTNASTLPQDIOEVGVYLIANQVQVPLRLRIVRSTQPLFEDNYALAVLDNG 120
61 ELTLPTNASTLPQDIOEVGVYLIANQVQVPLRLRIVRSTQPLFEDNYALAVLDNG 120
121 DPLANTPTVGTASFGGLRELOLRSLTBILKGGVLIQRNPOLCYODTILMKDIFRKNOLA 180
121 DPLANTPTVGTASFGGLRELOLRSLTBILKGGVLIQRNPOLCYODTILMKDIFRKNOLA 180
121 DPLANTPTVGTASFGGLRELOLRSLTBILKGGVLIQRNPOLCYODTILMKDIFRKNOLA 180
181 LTLIDTRSRACHPGKSGKSSGSSBDQSLTRVVCAGGACRCKGPLPTDCCHQC 240
181 LTLIDTRSRACHPGKSGKSSGSSBDQSLTRVVCAGGACRCKGPLPTDCCHQC 240
181 LTLIDTRSRACHPGKSGKSSGSSBDQSLTRVVCAGGACRCKGPLPTDCCHQC 240
241 AAGCTGPKHSDCLALCFPHNSGICELHCPALVTNTDTFESMPNPEGRTYFGASCVTACP 300
241 AAGCTGPKHSDCLALCFPHNSGICELHCPALVTNTDTFESMPNPEGRTYFGASCVTACP 300
241 AAGCTGPKHSDCLALCFPHNSGICELHCPALVTNTDTFESMPNPEGRTYFGASCVTACP 300
301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKSKPCARVCGVGLMEHLREAVATYSAN 360
301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKSKPCARVCGVGLMEHLREAVATYSAN 360
301 YNYLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKSKPCARVCGVGLMEHLREAVATYSAN 360
361 IOBPAGCKKTRFGLAPLPESFPDGPASNTAPLOPQOLVFETLBETITGYLIASMPDLP 420
361 IOBPAGCKKTRFGLAPLPESFPDGPASNTAPLOPQOLVFETLBETITGYLIASMPDLP 420
361 IOBPAGCKKTRFGLAPLPESFPDGPASNTAPLOPQOLVFETLBETITGYLIASMPDLP 420
421 DLAVFQNLQVTRGRILHNGAYSLTLOGLGISMLGRLSLREISGGLALHNTHTLCFVHTV 480
421 DLAVFQNLQVTRGRILHNGAYSLTLOGLGISMLGRLSLREISGGLALHNTHTLCFVHTV 480
421 DLAVFQNLQVTRGRILHNGAYSLTLOGLGISMLGRLSLREISGGLALHNTHTLCFVHTV 480
481 PMDQLFRNPQALHTANRPDECEVBEGLACHOLCARGHCWGPPTQCVCNSQFLRQEC 540
481 PMDQLFRNPQALHTANRPDECEVBEGLACHOLCARGHCWGPPTQCVCNSQFLRQEC 540
481 PMDQLFRNPQALHTANRPDECEVBEGLACHOLCARGHCWGPPTQCVCNSQFLRQEC 540
541 VBEGRVLOGLFRRVYNNAHCLPCHPCECPONGSVTCBEPZADQCAAHYDPPCVARC 600
541 VBEGRVLOGLFRRVYNNAHCLPCHPCECPONGSVTCBEPZADQCAAHYDPPCVARC 600
541 VBEGRVLOGLFRRVYNNAHCLPCHPCECPONGSVTCBEPZADQCAAHYDPPCVARC 600
601 PSQVAPDLSYMPIMKFPDEBGACOPCPINCHSCVDLDDKCPABORASPLTISAVVG 660
601 PSQVAPDLSYMPIMKFPDEBGACOPCPINCHSCVDLDDKCPABORASPLTISAVVG 660
601 PSQVAPDLSYMPIMKFPDEBGACOPCPINCHSCVDLDDKCPABORASPLTISAVVG 660

661 ILLVVLGVVFGILIRKROOKIRKTYMRLLQSTELVEPLTPSGAMPNOAMRLKXTEL 720
661 ILLVVLGVVFGILIRKROOKIRKTYMRLLQSTELVEPLTPSGAMPNOAMRLKXTEL 720
661 ILLVVLGVVFGILIRKROOKIRKTYMRLLQSTELVEPLTPSGAMPNOAMRLKXTEL 720
721 RAKVVGSGAFGVYKGIWIPDENYKIPVAIKYLRNTPSPKANKELIDEAYVAGVSP 780
721 RAKVVGSGAFGVYKGIWIPDENYKIPVAIKYLRNTPSPKANKELIDEAYVAGVSP 780
721 RAKVVGSGAFGVYKGIWIPDENYKIPVAIKYLRNTPSPKANKELIDEAYVAGVSP 780
781 YSRRLIGICTSTVQVLTQMPYGCILLDHVRENKRGSLGODLLNWCQIAKMSYLEDVR 840
781 YSRRLIGICTSTVQVLTQMPYGCILLDHVRENKRGSLGODLLNWCQIAKMSYLEDVR 840
781 YSRRLIGICTSTVQVLTQMPYGCILLDHVRENKRGSLGODLLNWCQIAKMSYLEDVR 840
841 LVHRDLAARVLYKSPNHVKITPGLARLLDIDETETHADGKRPICMMALESILRRRT 900
841 LVHRDLAARVLYKSPNHVKITPGLARLLDIDETETHADGKRPICMMALESILRRRT 900
841 LVHRDLAARVLYKSPNHVKITPGLARLLDIDETETHADGKRPICMMALESILRRRT 900
901 HQSDVMSYGVTVMLMTFGAKPYDGIPIAREIPDLLEKGERLPPOPCTIDVYIMVCKM 960
901 HQSDVMSYGVTVMLMTFGAKPYDGIPIAREIPDLLEKGERLPPOPCTIDVYIMVCKM 960
901 HQSDVMSYGVTVMLMTFGAKPYDGIPIAREIPDLLEKGERLPPOPCTIDVYIMVCKM 960
961 IDSECRPRFRELVSFBSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDMDGLVDA 1020
961 IDSECRPRFRELVSFBSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDMDGLVDA 1020
961 IDSECRPRFRELVSFBSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDMDGLVDA 1020
1021 EBYLVPOQGFPCDDPAPGAGMHHRRSSSTRSGGDLTLGLEPSEEARSPAPSEG 1080
1021 EBYLVPOQGFPCDDPAPGAGMHHRRSSSTRSGGDLTLGLEPSEEARSPAPSEG 1080
1021 EBYLVPOQGFPCDDPAPGAGMHHRRSSSTRSGGDLTLGLEPSEEARSPAPSEG 1080
1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLORYSEDPVPLPSETDGVYAPLTCSQPQREYV 1140
1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLORYSEDPVPLPSETDGVYAPLTCSQPQREYV 1140
1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLORYSEDPVPLPSETDGVYAPLTCSQPQREYV 1140
1141 NQPDVPRQPPSPREGPLPAARPGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
1141 NQPDVPRQPPSPREGPLPAARPGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
1141 NQPDVPRQPPSPREGPLPAARPGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
1201 GGAAPQHPPPAPSPARDNLYYWDQDPPERGAPSTKGPPTANPEYLGIDVVP 1255
1201 GGAAPQHPPPAPSPARDNLYYWDQDPPERGAPSTKGPPTANPEYLGIDVVP 1255
1201 GGAAPQHPPPAPSPARDNLYYWDQDPPERGAPSTKGPPTANPEYLGIDVVP 1255

RESULT 10
AAU77114 standard; Protein; 1255 AA.
AAU77114;
05-JUN-2002 (first entry)
Human Her-2/neu polypeptide.
Human; Her-2/neu; cytosolic; haematological malignancy; CML;
acute myelogenous leukaemia; AML; chronic myelogenous leukaemia; CLL;
chronic lymphocytic leukaemia; myeloma; non-Hodgkin's lymphoma; MDS;
Hodgkin's lymphoma; T cell therapy.
Homo sapiens.
MO200213847-A2.
21-FEB-2002.
13-AUG-2001; 2001WO-US25408.
14-AUG-2000; 2000US-0638280.
28-SEP-2000; 2000US-0675904.
(CORI-) CORIYA CORP.
Gaiger A, Cheever MA, Hand-zimmermann S;
WPI; 2002-280741/32.

DR N-PSDB, ABK10730.
 XX Inhibiting haematological malignancy development by administering
 PT polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide
 PT encoding the polypeptide, or antigen presenting cells expressing the
 PT polypeptide
 XX Disclosure, Page 71-74; 74pp; English.
 PS The invention relates to a method for inhibiting development of
 CC haematological malignancy in a patient by administering a polypeptide
 CC comprising an immunogenic portion of Her-2/neu or a polynucleotide
 CC encoding the polypeptide. Antigen presenting cells that express the
 CC protein can also be administered. The sequences are used for inhibiting
 CC development of haematological malignancy such as acute myelogenous
 CC leukemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic
 CC leukemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's
 CC lymphoma. This sequence represents the human Her-2/neu polypeptide.
 CC
 SQ Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 23; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIALACRMGLIALLPAGAASITQVCTGDMKRLPASPEITHLMDLRLHYOGCQVQGNL 60
 DB 1 MEIALACRMGLIALLPAGAASITQVCTGDMKRLPASPEITHLMDLRLHYOGCQVQGNL 60
 QY 61 ELTYLPTNALSISLQIQVQVYLIAHNOVROVPLORLAIYNGTQLFENYALAVLDNG 120
 DB 61 ELTYLPTNALSISLQIQVQVYLIAHNOVROVPLORLAIYNGTQLFENYALAVLDNG 120
 QY 121 DPLANTTPVTGASPGGLRELQSLRSLTEILKGVLIOBNPOLCYODITIMKDIFFHKNQLA 180
 DB 121 DPLANTTPVTGASPGGLRELQSLRSLTEILKGVLIOBNPOLCYODITIMKDIFFHKNQLA 180
 QY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDQSLTETVACGACGACRCKPLPTDCCHEQC 240
 DB 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDQSLTETVACGACGACRCKPLPTDCCHEQC 240
 QY 241 AAGCTGPKASDCLACHFNHSGICEHLCPALVTYNTDTPESMNPBGRYTFGASCTTACP 300
 DB 241 AAGCTGPKASDCLACHFNHSGICEHLCPALVTYNTDTPESMNPBGRYTFGASCTTACP 300
 QY 301 YNYLSTDVSGCTVCPLANOEVTABDGTORCEKSKPCAVCYGLGMEHLREVAATVSAN 360
 DB 301 YNYLSTDVSGCTVCPLANOEVTABDGTORCEKSKPCAVCYGLGMEHLREVAATVSAN 360
 QY 361 IOEPACCKKI FGSIAFLPESGFDGDPASNTAPLQPEOLQVETLEITGYLIISAMPDLP 420
 DB 361 IOEPACCKKI FGSIAFLPESGFDGDPASNTAPLQPEOLQVETLEITGYLIISAMPDLP 420
 QY 421 DLSVFONLOVIRGRIILHNGAYSLTLOGLGISWLSLSLRELSGGLALIHNTHLCPVHTV 480
 DB 421 DLSVFONLOVIRGRIILHNGAYSLTLOGLGISWLSLSLRELSGGLALIHNTHLCPVHTV 480
 QY 481 PMDQLFRNPHQALLHTANREDECVBEGIALCHQLCARGHWGSGPPQCVNCSQFLGQEC 540
 DB 481 PMDQLFRNPHQALLHTANREDECVBEGIALCHQLCARGHWGSGPPQCVNCSQFLGQEC 540
 QY 541 VEECRVLOGLPREYVNAHCLPCGPECOFONGSVTCGFPADQCVACAHYKDPFCVARC 600
 DB 541 VEECRVLOGLPREYVNAHCLPCGPECOFONGSVTCGFPADQCVACAHYKDPFCVARC 600
 QY 601 PSGVKEPDLSTMPYWKTPDEEGACQPCFINCTHS CVLDLDDGCPAEGRASPLTISIISAVG 660
 DB 601 PSGVKEPDLSTMPYWKTPDEEGACQPCFINCTHS CVLDLDDGCPAEGRASPLTISIISAVG 660
 QY 661 ILLVVLGVVFGILLIRROOKIRKTYMRRLLOETELVBLTPSGAMPNQAOMRILKETEL 720
 DB 661 ILLVVLGVVFGILLIRROOKIRKTYMRRLLOETELVBLTPSGAMPNQAOMRILKETEL 720

QY 721 RKVKVLSGAGFTVYKGIWIPDGENVKI PVAIKVLAENTSPKANKELIDEAYVAGVSGP 780
 DB 721 RKVKVLSGAGFTVYKGIWIPDGENVKI PVAIKVLAENTSPKANKELIDEAYVAGVSGP 780
 QY 781 YSRLLIGICITSTVOLVTOAMPGLCLDHYRENRGRISQDILLWCNQIAGMSYLEVR 840
 DB 781 YSRLLIGICITSTVOLVTOAMPGLCLDHYRENRGRISQDILLWCNQIAGMSYLEVR 840
 QY 841 LVHRDLAARNVLYKSPNHVKITPGLARLLDIDETEHADGKVPIMMALESILRRFT 900
 DB 841 LVHRDLAARNVLYKSPNHVKITPGLARLLDIDETEHADGKVPIMMALESILRRFT 900
 QY 901 HQSDVMSYGVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWVKCM 960
 DB 901 HQSDVMSYGVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWVKCM 960
 QY 961 IDSECPREPELYSESRMARDQRFVYVIONEDLGPSPLDSTYRSLLBEDDDGLVDA 1020
 DB 961 IDSECPREPELYSESRMARDQRFVYVIONEDLGPSPLDSTYRSLLBEDDDGLVDA 1020
 QY 1021 EETLVPOQGFPCDPAAGAMVHHRSSSTRSGGDLTLGLEPSESEAPRSLAPSEG 1080
 DB 1021 EETLVPOQGFPCDPAAGAMVHHRSSSTRSGGDLTLGLEPSESEAPRSLAPSEG 1080
 QY 1081 AGSDVFDGDLGMAAGLQSLPTHDPSPLOQYSEDPVLPSETDGYVAPLTCSPOPEYV 1140
 DB 1081 AGSDVFDGDLGMAAGLQSLPTHDPSPLOQYSEDPVLPSETDGYVAPLTCSPOPEYV 1140
 QY 1141 NOPDVAPPPSPREGLPAPARPAATLBERPKTISPGKNGVYKQVFAAGAVENBEYLTPQ 1200
 DB 1141 NOPDVAPPPSPREGLPAPARPAATLBERPKTISPGKNGVYKQVFAAGAVENBEYLTPQ 1200
 QY 1201 GGAAPQHPPPAPAFDNLXYMDOPPERGAPSPSTKGTPTAENPEYLGLDVAV 1255
 DB 1201 GGAAPQHPPPAPAFDNLXYMDOPPERGAPSPSTKGTPTAENPEYLGLDVAV 1255

RESULT 11
 AA92620
 ID AA92620 standard; Protein; 1255 AA.
 XX
 AC AA92620;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Human heregulin 2 (Her2).
 XX
 KW Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
 KM self-protein; cancer; breast cancer; prostate cancer;
 XX cell-associated peptide antigen; foreign epitope.
 OS Homo sapiens.
 XX
 FH Location/Qualifiers
 FT 1..173
 FT /label= N-terminal
 FT /note= "mature polypeptide"
 FT 5..25
 FT /label= insertion_region
 FT /note= "suitable for foreign epitope insertion"
 FT 59..73
 FT /label= insertion_region
 FT /note= "suitable for foreign epitope insertion"
 FT 103..117
 FT /label= insertion_region
 FT /note= "suitable for foreign epitope insertion"
 FT 149..163
 FT /label= insertion_region
 FT /note= "suitable for foreign epitope insertion"
 FT 174..323
 FT /label= Cysteine_rich_domain
 FT 210..224
 FT /label= insertion_region

PT		/note= "suitable for foreign epitope insertion"
FT	Region	250..264
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Domain	324..483
FT		/label= ligand_binding_domain
FT	Region	325..339
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Region	369..383
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Region	465..479
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Domain	484..623
FT		/label= Cysteine_rich_domain
FT	Region	579..593
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Domain	624..654
FT		/label= Transmembrane_domain
FT	Region	632..652
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Region	653..667
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Domain	655..1010
FT		/label= Tyrosine_kinase_domain
FT	Region	661..675
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Region	695..709
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Region	710..730
FT		/label= insertion_region
FT		/note= "suitable for foreign epitope insertion"
FT	Domain	1011..1235
FT		/label= C-terminal_domain
XX		
PN	WO200020027-A2.	
PD	13-APR-2000.	
XX		
PE	05-OCT-1999;	99WO-DK00525.
XX		
PR	05-OCT-1998;	98DK-0001261.
PR	20-OCT-1998;	98US-0105011.
XX		
PA	(MEBI-) M & B BIOTECH AS.	
XX		
PI	Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I,	
P1	Gautam A, Birk P, Karlsson G;	
XX		
DR	MP1; 2000-349917/30.	
DR	N-PsDB; AAA09455.	
XX		
PT	Inducing immune responses to weakly immunogenic, tumor associated	
PT	peptide antigens for the treatment of breast and prostate cancer	
XX		
PS	Claim 62; Page 193-198; 22pp; English.	
XX		
CC	This is the human hergulin 2 (Her2) sequence. Immunogenic analogues of	
CC	Her2 can be used in the claimed method as an autovaccine to induce a CTL	
CC	response. Subdominant CTL epitopes, antibody binding regions and	
CC	cysteine residues involved in disulfide bonds are preserved in the	
CC	immunogenized forms. Regions suitable for the insertion of foreign T	
CC	helper epitopes were identified (see features table). The method	
CC	is used for inducing immune responses against weakly immunogenic	
CC	cell-associated peptide antigens (PA) such as those associated with	
CC	cancers (self-proteins), e.g. human prostate specific membrane antigen	

CC	(PSM), hetergulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b).
CC	The method comprises effecting simultaneous presentation by antigen
CC	producing cells (APCs) of the animals immune system of: (1) at least 1
CC	CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC	B-cell group derived from the cell-associated PA; and (2) at least 1
CC	first T helper cell group which is foreign to the animal. Analogues of
CC	human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC	part of all known and predicted CTL and B-cell epitopes of the respective
CC	PA and including at least one foreign T helper epitope are also claimed.
CC	The method is used to treat prostate, prostate/breast or breast cancer
CC	when the PA is human PSM, FGF8b and Her2, respectively.
XX	
SQ	Sequence 1255 AA;
Query Match	99.9%; Score 6806; DB 21; Length 1255;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 1253;	Conservative 1; Mismatches 1; Indels 0; Gaps 0
QY	1 MELAALCRWGLLALPPGAASVQCTGTMKRLPASPETHLDMRLHYGCGQVVGNTL 60
DB	1 MELAALCRWGLLALPPGAASVQCTGTMKRLPASPETHLDMRLHYGCGQVVGNTL 60
QY	61 ELTYLPTNASSFLQDIOEVGVYLIANOVROVLOGLRIYRGVQLPEDNALAVLDNG 120
DB	61 ELTYLPTNASSFLQDIOEVGVYLIANOVROVLOGLRIYRGVQLPEDNALAVLDNG 120
QY	121 DELNNTTQVYTGASPGGLRELQRLSLTELKGGVLIQRNPOLCYOPTILMKDIFPKNNOLA 180
DB	121 DELNNTTQVYTGASPGGLRELQRLSLTELKGGVLIQRNPOLCYOPTILMKDIFPKNNOLA 180
QY	181 LTLIDTNRSRACHPCSPMCKSRGSSSEDCQSILTRTVCAAGCARCKGRLPTDCHEQC 240
DB	181 LTLIDTNRSRACHPCSPMCKSRGSSSEDCQSILTRTVCAAGCARCKGRLPTDCHEQC 240
QY	241 AAGCGPRGHSQCLACLFHNSHSGICELHCPALVYTYTDFPESNPNEGTYTPASCVTACP 300
DB	241 AAGCGPRGHSQCLACLFHNSHSGICELHCPALVYTYTDFPESNPNEGTYTPASCVTACP 300
QY	301 YNYLSTDVGSCTLYCPFLHNOEVTABDGTQRCESKSPCARVCYVGMEHLREVRATVSAN 360
DB	301 YNYLSTDVGSCTLYCPFLHNOEVTABDGTQRCESKSPCARVCYVGMEHLREVRATVSAN 360
QY	361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLQEPOLQVFETLEIITGYLYISAMPDSLIP 420
DB	361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLQEPOLQVFETLEIITGYLYISAMPDSLIP 420
QY	421 DLSPVQNIQVRRGLIHNGAVSLTIOGIGISMLGIRSLRELSGALLIHHNHLCEVHTV 480
DB	421 DLSPVQNIQVRRGLIHNGAVSLTIOGIGISMLGIRSLRELSGALLIHHNHLCEVHTV 480
QY	481 PMDOLFRRPHQALLHTANRPEDECVGEGLACGOLCARHGCMGPGFTQCVNCSQFLRGDEC 540
DB	481 PMDOLFRRPHQALLHTANRPEDECVGEGLACGOLCARHGCMGPGFTQCVNCSQFLRGDEC 540
QY	541 VEECRVLOGLPREYVNAHCLPCHECOPONGSVTCFSPBADQVACAHAHYKDPFCVARC 600
DB	541 VEECRVLOGLPREYVNAHCLPCHECOPONGSVTCFSPBADQVACAHAHYKDPFCVARC 600
QY	601 PEGVAPDLSYMPIMKFPDEBEGACOPCPINCHSCVDLDDKGPAPQARSPLTISIYAVG 660
DB	601 PEGVAPDLSYMPIMKFPDEBEGACOPCPINCHSCVDLDDKGPAPQARSPLTISIYAVG 660
QY	661 ILVVVGLVNGVIGIILKRQOKIRKTYMRLLQETLVEPLPSPGAMPNOAKRIKETEL 720
DB	661 ILVVVGLVNGVIGIILKRQOKIRKTYMRLLQETLVEPLPSPGAMPNOAKRIKETEL 720
QY	721 RKVKVLGSGAGTGYKGIWIPDGENVKIPVAIKVLRENTSPKANEILDEAYVMGVGSP 780
DB	721 RKVKVLGSGAGTGYKGIWIPDGENVKIPVAIKVLRENTSPKANEILDEAYVMGVGSP 780
QY	781 YVSRLLGICLTSTYQVLTQAMPYGLDHYENRGLSSQDILNMCQIAKMSYLEDYR 840
DB	781 YVSRLLGICLTSTYQVLTQAMPYGLDHYENRGLSSQDILNMCQIAKMSYLEDYR 840

QY 841 LVHRDLAARVAVLVKSNHVKITDPGLARLIDIDETVHADGKVPKIMMALISILRRRT 900
 DB 841 LVHRDLAARVAVLVKSNHVKITDPGLARLIDIDETVHADGKVPKIMMALISILRRRT 900
 QY 901 HOSDVSXVAVTWELMTFGAKPYDGI PAREIPOLLKGBRLPQRPICITDVMIMKCM 960
 DB 901 HOSDVSXVAVTWELMTFGAKPYDGI PAREIPOLLKGBRLPQRPICITDVMIMKCM 960
 QY 961 ISECRPRRRLVSESRMARADPQRFVIONEDLGPASPLDSTFYSLIEDDMGDLVDA 1020
 DB 961 ISECRPRRRLVSESRMARADPQRFVIONEDLGPASPLDSTFYSLIEDDMGDLVDA 1020
 QY 1021 EBYLVQOGFPCPDPAAGAGVHHRRSSSTSGGGLTLGLSPSEEAAPSPLAPSEG 1080
 DB 1021 EBYLVQOGFPCPDPAAGAGVHHRRSSSTSGGGLTLGLSPSEEAAPSPLAPSEG 1080
 QY 1081 AGSDVDFDGLGMAAGLQSLPTHDPSPLOVSEDPVPLPSTTDGVNAPLTCSPQPEVY 1140
 DB 1081 AGSDVDFDGLGMAAGLQSLPTHDPSPLOVSEDPVPLPSTTDGVNAPLTCSPQPEVY 1140
 QY 1141 NOPDVAPQPPSPREGPLPAARPAATLERPKTSLPGKNGVAVDVPAGAVENPEYLTPQ 1200
 DB 1141 NOPDVAPQPPSPREGPLPAARPAATLERPKTSLPGKNGVAVDVPAGAVENPEYLTPQ 1200
 QY 1201 GGAAPQHPAPSPAPADNLVYWDQDPERGAPSTFKGTPTAENPEYGLADV 1255
 DB 1201 GGAAPQHPAPSPAPADNLVYWDQDPERGAPSTFKGTPTAENPEYGLADV 1255

RESULT 12

AAE12130
 ID AAE12130 standard; Protein; 1255 AA.

AAE12130;

18-DEC-2001 (first entry)

Human tyrosine kinase-type receptor, HER-2.

Therapeutic compound: major histocompatibility complex; vaccine;
 antigenic peptide; MHC; immunoregulatory; immune response; HER-2;
 adoptive immunotherapy; anti-cancer; breast cancer antigen; APC;
 antigen presenting cell; human; tyrosine kinase-type receptor.

Homo sapiens.

Key Location/Qualifiers
 Region 774..782
 /note="Antigenic epitope"

WO200168677-A2.

20-SEP-2001.

16-MAR-2001; 2001WO-US40328.

16-MAR-2000; 2000US-0527487.

(GENZ) GENZYME CORP.

Nicolette CA;

WPI; 2001-616284/71.

N-PSDB; AAD19731.

Novel synthetic therapeutic compound for inducing immune response and
 for use in adoptive immunotherapy, has enhanced binding to major
 histocompatibility molecules and enhanced immunoregulatory properties

Claim 4; Page 63-67; 69pp; English.

CC The invention relates to synthetic therapeutic compounds (antigenic
 CC peptides) with enhanced binding to major histocompatibility complex
 CC (MHC) molecules and enhanced immunoregulatory properties relative
 CC to their natural counterparts. Compounds of the invention are useful
 CC for inducing an immune response in a subject and for use in adoptive
 CC immunotherapy. They are useful as components of anti-cancer vaccines
 CC and to expand immune effector cells that are specific for cancers
 CC characterized by expression of the breast cancer antigen, HER-2.
 CC Polynucleotides that encode peptides of the invention are useful as
 CC hybridization probes and as primers for the detection of genes of gene
 CC transcripts that are expressed in antigen presenting cells (APCs), to
 CC confirm transduction of polynucleotides into host cells. The present
 CC sequence is human tyrosine kinase-type receptor, HER-2. Compounds
 CC of the invention are designed based on the HER-2 antigenic peptide
 CC (774-782).

SO Sequence 1255 AA;

Query Match 99.9%; Score 6806; DB 22; Length 1255;

Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MELAALCRKGLLALIPGAASVOVCTGDMKRLRLASPEITHDMLRHLYOGGQVVGNL 60
 DB 1 MELAALCRKGLLALIPGAASVOVCTGDMKRLRLASPEITHDMLRHLYOGGQVVGNL 60
 QY 61 ELTYLPTNASLFLQDIQEVQVYLAHNOVQVPLQRLIYAGTOLFEDNVALAVLDNG 120
 DB 61 ELTYLPTNASLFLQDIQEVQVYLAHNOVQVPLQRLIYAGTOLFEDNVALAVLDNG 120
 QY 121 DELNNTPTVAGSPGGLRELQRLSTELLKGVLIQRPOLCYQDTITLMDIHKNNOLA 180
 DB 121 DELNNTPTVAGSPGGLRELQRLSTELLKGVLIQRPOLCYQDTITLMDIHKNNOLA 180
 QY 181 LTLIDNBRACHPCSPMKSGSRCSSESDQSLTRTYCAGGACAKGGLPDDCCEOC 240
 DB 181 LTLIDNBRACHPCSPMKSGSRCSSESDQSLTRTYCAGGACAKGGLPDDCCEOC 240
 QY 241 AAGCTGPKSDCLACLFPHNSGICELHCPALVTYNTDFESMNPREGRYTFGASCVTACP 300
 DB 241 AAGCTGPKSDCLACLFPHNSGICELHCPALVTYNTDFESMNPREGRYTFGASCVTACP 300
 QY 301 YNVLSTDVSCCTLVCELAHQEVTAEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
 DB 301 YNVLSTDVSCCTLVCELAHQEVTAEDGTORCEKSKPCARVCYGLGMEHLREYRAVTSAN 360
 QY 361 IQEPAGCKKI FGSGLAFLPESFGDDPASNTAPLOPBOLOVETLEETGYLYISAMPDLP 420
 DB 361 IQEPAGCKKI FGSGLAFLPESFGDDPASNTAPLOPBOLOVETLEETGYLYISAMPDLP 420
 QY 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWGLSLRELGSGLLIHNTHLCPHTV 480
 DB 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWGLSLRELGSGLLIHNTHLCPHTV 480
 QY 481 PMDQLFRNPHQALLHTANRPEDECVBEGLAHQLCARGCMGPTQCVNCSQFLRGQC 540
 DB 481 PMDQLFRNPHQALLHTANRPEDECVBEGLAHQLCARGCMGPTQCVNCSQFLRGQC 540
 QY 541 VEECRVLOGLPRBYNARHCLPCHPECOPONGSVTCFGEADOCVACAHYKDPFCVARC 600
 DB 541 VEECRVLOGLPRBYNARHCLPCHPECOPONGSVTCFGEADOCVACAHYKDPFCVARC 600
 QY 601 PSQVKPDLSTYMPIMKCPDEEGACQPCPINCSTSCVLDLDPGCPAEOBASPLTISAVVG 660
 DB 601 PSQVKPDLSTYMPIMKCPDEEGACQPCPINCSTSCVLDLDPGCPAEOBASPLTISAVVG 660
 QY 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOQMLIKETEL 720
 DB 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOQMLIKETEL 720
 QY 721 RKVKVLGSGAGFTVYKGIWIPGSENVKIPVALKVLRENTSPKANKELIDAYYAGVGP 780
 DB 721 RKVKVLGSGAGFTVYKGIWIPGSENVKIPVALKVLRENTSPKANKELIDAYYAGVGP 780

QY 781 YSRRLGICLTSTVQVLTQMLPVCCLLDHYRENRGRLGSODLLNMCQIAKMSYLEDVR 840
 Db 781 YSRRLGICLTSTVQVLTQMLPVCCLLDHYRENRGRLGSODLLNMCQIAKMSYLEDVR 840
 QY 841 LVHRDLAARNVLYKSPNNHKITDFGLARLLDIDETETHAOGKVPKIMMALESTLRRT 900
 Db 841 LVHRDLAARNVLYKSPNNHKITDFGLARLLDIDETETHAOGKVPKIMMALESTLRRT 900
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPPOPICITIDVYIMVCKM 960
 Db 901 HQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPPOPICITIDVYIMVCKM 960
 QY 961 IDSECRFRFRELVESESRMARDPORFVITQNEDELGPASPLDSTYRSLLEDDMDGLVDA 1020
 Db 961 IDSECRFRFRELVESESRMARDPORFVITQNEDELGPASPLDSTYRSLLEDDMDGLVDA 1020
 QY 1021 EETLVPOQGFPCPDPAAGAGMVAHHRSSSTRSGGDLTLGLEPSREBA PRSPLAPSEG 1080
 Db 1021 EETLVPOQGFPCPDPAAGAGMVAHHRSSSTRSGGDLTLGLEPSREBA PRSPLAPSEG 1080
 QY 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSEDPVPLPSETDGYAAPLTCSPQPEYV 1140
 Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSEDPVPLPSETDGYAAPLTCSPQPEYV 1140
 QY 1141 NQPDVPRQPPSPRPGPLPAAPRPAATLERPKTSPGKNGVYKOVFAFGAVENPEYLTPO 1200
 Db 1141 NQPDVPRQPPSPRPGPLPAAPRPAATLERPKTSPGKNGVYKOVFAFGAVENPEYLTPO 1200
 QY 1201 GGAAPQHPPPAPSPAPDNLTYMDQDPERGAPSTFKGTPTAENPYLGLADVAV 1255
 Db 1201 GGAAPQHPPPAPSPAPDNLTYMDQDPERGAPSTFKGTPTAENPYLGLADVAV 1255

RESULT 13
 AAB60167 standard; Protein; 1255 AA.

ID AAB60167 standard; Protein; 1255 AA.
 AC AAB60167;
 XX
 DT 03-APR-2001 (first entry)
 XX
 DE HER2 transgene plasmid construct encoded protein.
 XX
 KW Human; HER2; ErbB2 receptor; p185neu; maytansinoid conjugate; cancer;
 XX antibody.
 OS Homo sapiens.
 OS Synthetic.
 PN MO200100244-A2.
 PD 04-JAN-2001.
 PF 23-JUN-2000; 2000MO-US17229.
 PR 25-JUN-1999; 99US-0141316.
 PR 16-MAR-2000; 2000US-0189844.
 PA (GETH) GENENTECH INC.
 PI Erickson S, Schwall R;
 DR WPI; 2001-061962/07.
 DR N-PSDB; AAF24297.
 XX
 PT Treating tumors, particularly breast cancers, which overexpress an ErbB
 PT receptor and does not respond to an anti-ErbB antibody, comprises
 PT conjugating the antibody to a maytansinoid -
 PS
 PS Example 3; Fig 4; 92pp; English.
 CC The present invention provides a method of treating cancer by

CC administering a conjugate of anti-ErbB antibody with a maytansinoid. In
 CC particular, the antibody is directed against ErbB2 (also known as HER2
 CC and p185neu). The method is particularly useful in the treatment of
 CC breast, ovarian, stomach, endometrial, salivary gland, lung, kidney,
 CC colon, colorectal, thyroid, pancreatic, prostate and bladder cancers.

SQ Sequence 1255 AA;

Query Match 99.9%; Score 6806; DB 22; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 METALACRMGLIALLPAGASTQVCTGDMKRLPLASPTTHDMLRHLYOGCOVQGNL 60
 Db 1 METALACRMGLIALLPAGASTQVCTGDMKRLPLASPTTHDMLRHLYOGCOVQGNL 60
 QY 61 EETLVPTNASTSLFLODIQEVQGVVLAHNOVROVPLQRLRIVRGTOLEFENVYALAVDNG 120
 Db 61 EETLVPTNASTSLFLODIQEVQGVVLAHNOVROVPLQRLRIVRGTOLEFENVYALAVDNG 120
 QY 121 DELANTTPVTGASPGGLRELQRLSTELLKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
 Db 121 DELANTTPVTGASPGGLRELQRLSTELLKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
 QY 181 LHLIDTNRSPAHPSPCKSGRCWGBSSDCSLTRTVACAGCARCKGPLPTDCHEQC 240
 Db 181 LHLIDTNRSPAHPSPCKSGRCWGBSSDCSLTRTVACAGCARCKGPLPTDCHEQC 240
 QY 241 ANGCTGPKHSDCLACHPNHSGLCELAHPALVYNTDFESMNPBERYTFGASCVTACP 300
 Db 241 ANGCTGPKHSDCLACHPNHSGLCELAHPALVYNTDFESMNPBERYTFGASCVTACP 300
 QY 301 YNLTSTDVSGCTLVCPLANQEVTAEDGTORCEKSKRCARVCYGLGMEHLREVAVTSAN 360
 Db 301 YNLTSTDVSGCTLVCPLANQEVTAEDGTORCEKSKRCARVCYGLGMEHLREVAVTSAN 360
 QY 361 IQEFACCKKIFGSLAFLPESFDGPAANTAPLOPQLOVETLEBEITGYIYSAMPDLP 420
 Db 361 IQEFACCKKIFGSLAFLPESFDGPAANTAPLOPQLOVETLEBEITGYIYSAMPDLP 420
 QY 421 DISVFONLOVYIRIRILHNGAVSLTLQIGISWGLSLRLSGSLALIHNTNLCFPHTV 480
 Db 421 DISVFONLOVYIRIRILHNGAVSLTLQIGISWGLSLRLSGSLALIHNTNLCFPHTV 480
 QY 481 PMDQLFRNPHQALHTANRPEDECVSGGLACHOLCARGHCWGPPTQCVNCSQFLRGQEC 540
 Db 481 PMDQLFRNPHQALHTANRPEDECVSGGLACHOLCARGHCWGPPTQCVNCSQFLRGQEC 540
 QY 541 VEECRVLOGLPRRYVNAHRLPCHPECOPONGSVTCGPBADOCVACAHYKDPFCVARC 600
 Db 541 VEECRVLOGLPRRYVNAHRLPCHPECOPONGSVTCGPBADOCVACAHYKDPFCVARC 600
 QY 601 PSGVXPDLSTYMPIMKPPDEEGACOPCPINCTHSCVDLDDKGPABEORASPLTISIVVG 660
 Db 601 PSGVXPDLSTYMPIMKPPDEEGACOPCPINCTHSCVDLDDKGPABEORASPLTISIVVG 660
 QY 661 ILLVVVLGVVFGILLIRROOKIRKTYMRLLQETELVEPLTPSGAMENQAMRLKETEL 720
 Db 661 ILLVVVLGVVFGILLIRROOKIRKTYMRLLQETELVEPLTPSGAMENQAMRLKETEL 720
 QY 721 RAYKVLGSGAFGTVYKGIWI PDGENYKIPVAILVLEBNTSPKANKETLDEAYVAVAGVSP 780
 Db 721 RAYKVLGSGAFGTVYKGIWI PDGENYKIPVAILVLEBNTSPKANKETLDEAYVAVAGVSP 780
 QY 781 YSRRLGICLTSTVQVLTQMLPVCCLLDHYRENRGRLGSODLLNMCQIAKMSYLEDVR 840
 Db 781 YSRRLGICLTSTVQVLTQMLPVCCLLDHYRENRGRLGSODLLNMCQIAKMSYLEDVR 840
 QY 841 LVHRDLAARNVLYKSPNNHKITDFGLARLLDIDETETHAOGKVPKIMMALESTLRRT 900
 Db 841 LVHRDLAARNVLYKSPNNHKITDFGLARLLDIDETETHAOGKVPKIMMALESTLRRT 900
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPPOPICITIDVYIMVCKM 960

```

Db      |||||
901 HOSDWSYGVWELMTFGAKPYDGI.PABEIPDLLEKGERLPQPPICTIVYIMVACMM 960
Qy      961 IDSECRPRRELVSSESRMARDFORFVVIQNEIDGPASPLDSFIFYRSLBEDDMDGLVDA 1020
Db      961 IDSECRPRRELVSSESRMARDFORFVVIQNEIDGPASPLDSFIFYRSLBEDDMDGLVDA 1020
Qy      1021 BEYLVPQOGFFCPDPAFAGAGMVHRRHSSTRSGGDLTLGLEPSEEA PRSPLAPSEG 1080
Db      1021 BEYLVPQOGFFCPDPAFAGAGMVHRRHSSTRSGGDLTLGLEPSEEA PRSPLAPSEG 1080
Qy      1081 AGSDVDGDLGMAAGLGLSLPTHDSPLQRYSEDPTVPLPSTTDGYVALTCSPOPEYV 1140
Db      1081 AGSDVDGDLGMAAGLGLSLPTHDSPLQRYSEDPTVPLPSTTDGYVALTCSPOPEYV 1140
Qy      1141 NOPDVAPQPPSPREGPLPAARPAATLERPKTSLPGKNGVVDVPAFAGAVENPEYTPQ 1200
Db      1141 NOPDVAPQPPSPREGPLPAARPAATLERPKTSLPGKNGVVDVPAFAGAVENPEYTPQ 1200
Qy      1201 GGAAPQHPPEPAPSPAFDNLVYWDOPPERGAPSTPTKGTPTAENPEYGLDVPV 1255
Db      1201 GGAAPQHPPEPAPSPAFDNLVYWDOPPERGAPSTPTKGTPTAENPEYGLDVPV 1255

RESULT 14
AAE26349 standard; Protein; 1255 AA.
XX      AAE26349;
XX      13-DEC-2002 (first entry)
XX      Human HER-2 protein.
XX      Transgenic animal; transgenic; mammary gland cell; HER2; tumour;
XX      cancer; therapy; apoptosis; cytostatic; human.
XX      Homo sapiens.
XX      US2002035736-A1.
XX      21-MAR-2002.
XX      16-MAR-2001; 2001US-0811115.
XX      16-MAR-2000; 2000US-189844P.
XX      (ERIC/) ERICKSON S.
XX      (KING/) KING K.
XX      (SCHW/) SCHWALL R.
XX      Erickson S, King K, Schwall R;
XX      WPI; 2002-401155/43.
XX      N-PSDB; AAD43934, AAD43935.
XX      New transgenic non-human mammal that produces detectable levels of a
XX      native human HER2 protein in its mammary gland cells, useful as tumor
XX      models for testing HER2-directed cancer therapies, and for identifying
XX      anticancer agents.
XX      Example 2; Page 26-29; 83pp; English.
XX      The invention relates to a transgenic non-human mammal that produces in
XX      its mammary gland cells detectable levels of a native human HER2 protein
XX      or its fragment. The transgenic animals are useful as tumour models for
XX      testing HER2-directed cancer therapies, and for identifying anticancer
XX      agents. The animals may also be used as source of cells which can be
XX      immortalised in culture, in screening for compounds that have potential
XX      as prophylactic or therapeutic treatments of diseases or disorders
XX      involving expression of HER2. The anti-cancer molecules are useful for
XX      inducing apoptosis or cell death of cancer cells. The present sequence
XX      is human HER-2 protein.

```

```

XX      SQ      Sequence      1255 AA;
XX      Query Match      99.9%; Score 6806; DB 23; Length 1255;
XX      Best Local Similarity 99.8%; Pred. No. 0;
XX      Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy      1 MELALACRGLLALLPFGAASQVCTGMDKRLPASPETHLDMLRLHYOGCVOVGNL 60
Db      1 MELALACRGLLALLPFGAASQVCTGMDKRLPASPETHLDMLRLHYOGCVOVGNL 60
Qy      61 EUTYLPFNASLSFLQDIQEVGVYLIANQVROVPLQRLIRVGTQLFEDNYALAVLDNG 120
Db      61 EUTYLPFNASLSFLQDIQEVGVYLIANQVROVPLQRLIRVGTQLFEDNYALAVLDNG 120
Qy      121 DPLNNTPTVAGSPGRLRELQBSLREILKGGTLIQNPOLCYODTILMDIFHKNNOLA 180
Db      121 DPLNNTPTVAGSPGRLRELQBSLREILKGGTLIQNPOLCYODTILMDIFHKNNOLA 180
Qy      181 LTLIDTNRARACHPCSPMCKSGRCWGESSEDCSLTRTVCAAGCARGKPLPTDCHEQC 240
Db      181 LTLIDTNRARACHPCSPMCKSGRCWGESSEDCSLTRTVCAAGCARGKPLPTDCHEQC 240
Qy      241 AAGCTGPKASDCLACHFNHSGICEIHCPLVYNTDTFESMPNDEGRYTFGASCVTACP 300
Db      241 AAGCTGPKASDCLACHFNHSGICEIHCPLVYNTDTFESMPNDEGRYTFGASCVTACP 300
Qy      301 YNYLSTDVSCCTLVCPILHNOVTAEDGTORCEKSRPCARVCYGLGMEHLREVAVTSAN 360
Db      301 YNYLSTDVSCCTLVCPILHNOVTAEDGTORCEKSRPCARVCYGLGMEHLREVAVTSAN 360
Qy      361 IDEFAGCKKI FGSIAFLPESFDGPASNTAPLQBPOLQVETLEITGYLIISAMPDSL 420
Db      361 IDEFAGCKKI FGSIAFLPESFDGPASNTAPLQBPOLQVETLEITGYLIISAMPDSL 420
Qy      421 DLSVFQNLQVIRGRILHNGAYSITLQGLGI SWLGRLSRLSGSLALIHNNTHLCFVHTV 480
Db      421 DLSVFQNLQVIRGRILHNGAYSITLQGLGI SWLGRLSRLSGSLALIHNNTHLCFVHTV 480
Qy      481 PMDQLFRNHQALHNTANRPEDECVBEGSLACHQLCARGHCKMGFPQCVNCSQFLROEC 540
Db      481 PMDQLFRNHQALHNTANRPEDECVBEGSLACHQLCARGHCKMGFPQCVNCSQFLROEC 540
Qy      541 VEECRVLOGLPREYVVARHCLPCHPECOPONGSVTCFEPADQCVACAHKDPFCVARG 600
Db      541 VEECRVLOGLPREYVVARHCLPCHPECOPONGSVTCFEPADQCVACAHKDPFCVARG 600
Qy      601 PSGVKEPDLSTYMPIMKEPDEBGAQPCPINCTHSCVDLDKGCPRABORASPLTISAVVG 660
Db      601 PSGVKEPDLSTYMPIMKEPDEBGAQPCPINCTHSCVDLDKGCPRABORASPLTISAVVG 660
Qy      661 ILLVVLGVVFGILLKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRLKETEL 720
Db      661 ILLVVLGVVFGILLKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRLKETEL 720
Qy      721 RMYKVLGSGAPGVYVGIWIPGSENVKIPVAKVLEBNTSPKANKELIDEAYVAGVGP 780
Db      721 RMYKVLGSGAPGVYVGIWIPGSENVKIPVAKVLEBNTSPKANKELIDEAYVAGVGP 780
Qy      781 YVSRLLGICLTSTVOLVTLQMPYGLLDHVRENKRGSLGSDLLNMCQIAXKGSYLEDDR 840
Db      781 YVSRLLGICLTSTVOLVTLQMPYGLLDHVRENKRGSLGSDLLNMCQIAXKGSYLEDDR 840
Qy      841 LVHRDLAANVLYKSPNNHYKIDFGELARLIDDETYHADGKVPYKMAALSSILRRRFT 900
Db      841 LVHRDLAANVLYKSPNNHYKIDFGELARLIDDETYHADGKVPYKMAALSSILRRRFT 900
Qy      901 HOSDWSYGVWELMTFGAKPYDGI.PABEIPDLLEKGERLPQPPICTIVYIMVACMM 960
Db      901 HOSDWSYGVWELMTFGAKPYDGI.PABEIPDLLEKGERLPQPPICTIVYIMVACMM 960
Qy      961 IDSECRPRRELVSSESRMARDFORFVVIQNEIDGPASPLDSFIFYRSLBEDDMDGLVDA 1020

```

Db 961 IDSECRPRRELVSERSRMARDPQRFVYIQLNEDLGPAASPLDSTFYNSLLEDDMDGLVDA 1020
Qy 1021 BEYLVPQGGFCFCDPAPGAGGVVHHRRSSSTSSGGGDLTLGLEPSEEARPSPLAPSEG 1080
Db 1021 BEYLVPQGGFCFCDPAPGAGGVVHHRRSSSTSSGGGDLTLGLEPSEEARPSPLAPSEG 1080
Qy 1081 AGSDVDVDGLGMAAGLGLSLPTHDPSPLOKRYSEDPTVPLPSTDTGVNAPLTCSPQPEYV 1140
Db 1081 AGSDVDVDGLGMAAGLGLSLPTHDPSPLOKRYSEDPTVPLPSTDTGVNAPLTCSPQPEYV 1140
Qy 1141 NOPDVAPQPPSPREGPLPAAPAGATLERPKTSPGKGVVXOVAPFAGAVENPEYLTPO 1200
Db 1141 NOPDVAPQPPSPREGPLPAAPAGATLERPKTSPGKGVVXOVAPFAGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPPEAPSPAFDNLVYWDQDPERGAPSTFKGPTPAENPEYLGADV 1255
Db 1201 GGAAPQHPPEAPSPAFDNLVYWDQDPERGAPSTFKGPTPAENPEYLGADV 1255
RESULT 15
AAE26366
ID AAE26366 standard; Protein; 1255 AA.
XX AAE26366;
AC AAE26366;
XX 13-DEC-2002 (first entry)
DE Human Her2 antigen.
XX Human: Immune response; T-helper cell epitope; chitosan; CTL response;
KM vaccine; prostate cancer; breast cancer; Her2 antigen; cytostatic;
XX Immunostimulant.
XX Homo sapiens.
OS
XX
FH Key location/Qualifiers
FT Peptide 1..23
FT Protein /label= signal_peptide
FT 24..1255
/note= "Mature human Her2 antigen"
XX MO200234287-A2.
XX 02-MAY-2002.
XX 26-OCT-2001; 2001WO-DK00705.
XX 27-OCT-2000; 2000DK-0001606.
XX 03-NOV-2000; 2000US-245166P.
XX 18-JUN-2001; 2001DK-0000936.
XX (PHAR-) PHARMEXA AS.
XX Beier AM, Gautam A, Mouritsen S;
XX WPI; 2002-463339/49.
XX N-PSDB; AAD43986.
XX
XX Inducing or enhancing an immune response against an antigen,
PT particularly cytotoxic T-lymphocyte responses, for treating or
PT ameliorating prostate or breast cancer, comprises administering the
PT antigen formulated with chitosan
XX
XX Disclosure; Page 91-95; 97pp; English.
XX
XX The invention relates to a method for inducing or enhancing an immune
XX response against a polypeptide antigen in an animal, including human.
XX The method comprises administering the polypeptide antigen or at least
XX one variant which includes at least one first T-helper cell epitope that
XX is foreign to the animal (foreign T cell epitope) and is formulated with
XX chitosan. The polypeptide antigen is weakly immunogenic or non-
XX immunogenic. The invention is used as vaccine. The chitosan and
XX polypeptide antigen or its variant are useful in the preparation of an

CC immunogenic composition for inducing or enhancing an immune response,
CC particularly CTL response, against the polypeptide or protein antigen.
CC The method for inducing or enhancing an immune response is useful in
CC treating or ameliorating cancer, e.g. prostate or breast cancer. The
CC present sequence is human Her2 antigen.
XX
SQ Sequence 1255 AA;
Query Match 99.9%; Score 6806; DB 23; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MELAAICRGGGLIALLPAGASTQVCTGDMKRLPASPEPTHDMRLHLYOGCOVQGNL 60
Db 1 MELAAICRGGGLIALLPAGASTQVCTGDMKRLPASPEPTHDMRLHLYOGCOVQGNL 60
Qy 61 ELTYLPTNASLSTLODIOEVQGVYLIANHOVRQVPLQRLIYVGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSTLODIOEVQGVYLIANHOVRQVPLQRLIYVGTQLFEDNYALAVLDNG 120
Qy 121 DELNNTPTVYTGASPGGLRELQRLSTLEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
Db 121 DELNNTPTVYTGASPGGLRELQRLSTLEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
Qy 181 LTLIDPNSRACHPGSPMCKGSRQWGSSESDCSLRTVCAGGACARCKGLPTDCCHEOC 240
Db 181 LTLIDPNSRACHPGSPMCKGSRQWGSSESDCSLRTVCAGGACARCKGLPTDCCHEOC 240
Qy 241 AGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTESMNPREGRYTFGASCVTACP 300
Db 241 AGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTESMNPREGRYTFGASCVTACP 300
Qy 301 YNTLSTDVSGCTLVCPLENOEVTABDGTQRCCKSPCARVCYGLGMEHLREVRATVSAN 360
Db 301 YNTLSTDVSGCTLVCPLENOEVTABDGTQRCCKSPCARVCYGLGMEHLREVRATVSAN 360
Qy 361 IOEFACCKKIFGSLALPESFDSDPSNTRPLQPEOLQVETLEITGVIYISMPSLP 420
Db 361 IOEFACCKKIFGSLALPESFDSDPSNTRPLQPEOLQVETLEITGVIYISMPSLP 420
Qy 421 DISVFNOLQVIRIRILHNGAYSLTLQGLISWGLSLRELGSGLAIHNTHLCFVHTV 480
Db 421 DISVFNOLQVIRIRILHNGAYSLTLQGLISWGLSLRELGSGLAIHNTHLCFVHTV 480
Qy 481 PMDQLFRNPHQALHTANRPEDECVBGLACHQLCARGHCWGPGPTCVNCSQPLRQEC 540
Db 481 PMDQLFRNPHQALHTANRPEDECVBGLACHQLCARGHCWGPGPTCVNCSQPLRQEC 540
Qy 541 VBECKVLOGLPREYVARHCLPCHPSCOPONGSVTCEGPEADOCVACAHYKDPFCVARC 600
Db 541 VBECKVLOGLPREYVARHCLPCHPSCOPONGSVTCEGPEADOCVACAHYKDPFCVARC 600
Qy 601 PGGKPEDLASYMPIKPEDEGACQPCPINCTHSCVDLDKGCFAEQAASPLTISAVVG 660
Db 601 PGGKPEDLASYMPIKPEDEGACQPCPINCTHSCVDLDKGCFAEQAASPLTISAVVG 660
Qy 661 ILTVVVLGVVFGILIRKROOKIRKTYMRRLQSTELVEPLTPSGAMNOMRLKTEL 720
Db 661 ILTVVVLGVVFGILIRKROOKIRKTYMRRLQSTELVEPLTPSGAMNOMRLKTEL 720
Qy 721 RKRYKVGSGAFGVYGIWIPDEGNVKIPAIKYLRENTSPKANKETLDEAYVYVAGVSP 780
Db 721 RKRYKVGSGAFGVYGIWIPDEGNVKIPAIKYLRENTSPKANKETLDEAYVYVAGVSP 780
Qy 781 YVSRLLGICITSTVQVLTQMPYGCILLDHVRENKRGSLDLNMCQIAKMSYLEDVR 840
Db 781 YVSRLLGICITSTVQVLTQMPYGCILLDHVRENKRGSLDLNMCQIAKMSYLEDVR 840
Qy 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDDETETHAGKVPKPMALSTLRRT 900
Db 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDDETETHAGKVPKPMALSTLRRT 900
Qy 901 HQSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGBRLPOPICTIDVYIMVNCWM 960

Db	901	HQSDVMSYGVTVWELMTFGAKFYDGIIPAREIPDLLEKGERLPDPPICTIDVTIMVKCM	960
Qy	961	IDSECRPRFRELVSSEFSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSILBEDDMGDLVDA	1020
Db	961	IDSECRPRFRELVSSEFSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSILBEDDMGDLVDA	1020
Qy	1021	EEYLVPOQGFPCDPAAGAGMHHRRSSSTRSGGDLTLGLEPSEEEAPRSPAPSEG	1080
Db	1021	EEYLVPOQGFPCDPAAGAGMHHRRSSSTRSGGDLTLGLEPSEEEAPRSPAPSEG	1080
Qy	1081	AGSDVDFGDLGMAKGLQSLPTHDPSPLORYSEDPVPLPSETDGYVAPLTCSPQBEYV	1140
Db	1081	AGSDVDFGDLGMAKGLQSLPTHDPSPLORYSEDPVPLPSETDGYVAPLTCSPQBEYV	1140
Qy	1141	NQPDVAPQPPSPREGPLPAARPAATLERPKTLPQNGVVKOVFAFGAVENPEYLTPO	1200
Db	1141	NQPDVAPQPPSPREGPLPAARPAATLERPKTLPQNGVVKOVFAFGAVENPEYLTPO	1200
Qy	1201	GGAAPQPHPPAPSPAFDNLTYWDQDPPERGAAPPSTFKGTPTAENPEYLGLDVAV	1255
Db	1201	GGAAPQPHPPAPSPAFDNLTYWDQDPPERGAAPPSTFKGTPTAENPEYLGLDVAV	1255

Search completed: February 9, 2004, 16:15:14
 Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: February 9, 2004, 16:13:07 ; Search time 27 Seconds

(without alignments)
4470.063 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815
Sequence: 1 MEALALCRWGLLLALPEGA.....TFKGTPTAENPEYLGLDPV 1255Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

PIR 76: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6815	100.0	1255	1 A24571	protein-tyrosine k
2	5997	88.0	1260	1 TYRTMU	protein-tyrosine k
3	5993.5	87.9	1254	1 I48161	p-185 precursor -
4	3169	46.5	1210	1 GQHUR	epidermal growth f
5	3145	46.1	1210	2 A53183	epidermal growth f
6	3122.5	45.8	1223	1 TYCHLV	epidermal growth f
7	3004.5	44.1	1308	2 A47253	epidermal growth f
8	2708	39.7	1166	1 S06142	protein-tyrosine k
9	2432.5	35.7	1342	2 A36223	kinase-related tra
10	2347.5	34.4	1339	2 UC4387	epidermal growth f
11	1765.5	25.9	698	1 TVFVIV	protein-tyrosine k
12	1702	25.0	604	1 TVYUH	epidermal growth f
13	1653.5	24.3	1330	1 GQFPE	epidermal growth f
14	1645	24.1	544	2 S35745	protein-tyrosine k
15	1638	24.0	545	2 S00727	kinase-related tra
16	1621	23.8	540	2 B44776	protein-tyrosine k
17	1619	22.5	540	2 TVPVB	protein-tyrosine k
18	1536	22.5	644	2 A36355	epidermal growth f
19	1301	19.1	1323	2 E88257	protein let-23 [lm
20	1301	19.1	1374	2 S70712	protein-tyrosine k
21	1211	17.8	1369	2 S70713	protein-tyrosine k
22	1177	17.3	1717	1 A45558	epidermal growth f
23	1155	16.9	527	2 A42031	epidermal growth f
24	997.5	14.6	843	2 A27131	epidermal growth f
25	814.5	12.0	346	2 S13807	protein-tyrosine k
26	757	11.1	311	2 S13808	protein-tyrosine k
27	736	10.8	1363	2 T43220	insulin-like growt
28	717	10.5	1362	1 INHUR	insulin receptor p
29	710	10.4	1363	2 A36080	insulin receptor p

30	709.5	10.4	1372	2 A34157	insulin receptor p
31	703.5	10.3	1607	2 T43212	insulin-like growt
32	693.5	10.2	1300	2 A36502	insulin receptor-r
33	682.5	10.0	1477	2 T18534	protein-tyrosine k
34	681	10.0	1268	2 B36502	insulin receptor-r
35	651	9.6	1367	1 IGHUR1	insulin-like growt
36	642	9.4	1371	2 A33837	insulin-like growt
37	627	9.2	1390	2 T30346	insulin receptor -
38	624.5	9.2	2148	1 A56081	insulin receptor -
39	622	9.1	2101	2 S57245	insulin receptor (
40	607	8.9	987	2 A54052	protein-tyrosine k
41	591	8.7	1091	2 S33596	protein-tyrosine k
42	590.5	8.7	952	2 I50612	protein-tyrosine k
43	589.5	8.7	977	2 S49004	tyrosine kinase Mp
44	587	8.6	1114	1 S05582	protein-tyrosine k
45	585	8.6	987	2 I48652	mouse developmenta

ALIGNMENTS

RESULT 1

A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N/Alternate names: C-erb-B-2 protein precursor; Kinase-related transforming protein erb
C/Species: Homo sapiens (man)
C/Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999
C/Accession: A24571, A25491, I44188, B44188, I59509, I57622
R/Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;
Nature 319, 230-234, 1986
A/Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth
A/Reference number: A24571, MUID:86118663, PMID:3303577
A/Accession: A24571
A/Molecule type: mRNA
A/Residues: 1-1255 <YAM>
A/Cross-references: GB:X03363, NID:G31197, PIDN:CAA27060.1, PID:G31198
R/Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A/Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epider
A/Reference number: A25491, MUID:86016723, PMID:2995967
A/Accession: A25491
A/Molecule type: DNA
A/Residues: 737-1031 <SEM>
A/Molecule type: DNA
A/Cross-references: GB:M11767, NID:G182163, PIDN:AAA35808.1, PID:G553282
R/Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P.
Science 230, 1132-1139, 1985
A/Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromo
A/Reference number: A44188, MUID:86070181, PMID:2999974
A/Accession: A44188
A/Molecule type: DNA
A/Residues: 740-910 <COU>
A/Cross-references: GB:M12036, NID:G183988, PIDN:AAA35978.1, PID:G183989
A/Accession: B44188
A/Molecule type: mRNA
A/Residues: 1-517, 'RAL', '522', 'S', '524-654', 'V', '656-1169', 'A', '1171-1255 <COU>
A/Cross-references: GB:M11730, NID:G183986
R/King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 228, 974-976, 1985
A/Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A/Reference number: I59509, MUID:85272597, PMID:2992089
A/Accession: I59509
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 832-909 <REX>
A/Cross-references: GB:L29395, NID:G459807, PIDN:AAA35809.1, PID:G459808
R/Tal, M.; King, C.R.; Kraus, M.H.; Ulrich, A.; Schleefinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A/Title: Human HBR2 (neu) promoter: evidence for multiple mechanisms for transcriptional
A/Reference number: I57622, MUID:87268898, PMID:3039351
A/Accession: I57622
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-191 <TML>

A/Cross-references: GB:M16792; NID:g183983; PIDN:AAA58637.1; PID:g553332
 C/Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
 C/Genetics:
 A:Gene: GDB:ERBB2; NGL; NEU; HER-2
 A/Cross-references: GDB:120613; OMIM:164870
 A/Map position: 17q21.1-17q21.1
 A/Introns: 25/1; 75/3; 147/1; 883/3
 A/Note: the list of introns is incomplete
 C/Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphinase
 F/1-21/Domain: signal sequence #status predicted <SIG>
 F/22-125/Produce: protein-tyrosine kinase erbB2 #status predicted <MAT>
 F/22-653/Domain: extracellular #status predicted <EXT>
 F/70-304/Domain: EGF receptor extracellular domain repeat <EB1>
 F/395-605/Domain: EGF receptor extracellular domain repeat <EB2>
 F/554-675/Domain: transmembrane #status predicted <TM>
 F/718-983/Domain: intracellular #status predicted <INT>
 F/726-734/Region: protein kinase ATP-binding motif
 F/68,124,187,259,530,571,629/Binding site: carbohydrate (asn) (covalent) #status predict
 F/686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F/753/Active site: Lys #status predicted
 F/1139,1221,1222,1246/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 100.0%; Score 6815; DB 1; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 1,1e-270;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEALALCRMGILLALLPRGAASVCTGTDMKRLPASPTHLDMLRHLVYGCQVVOGNTL 60
 DB 1 MEALALCRMGILLALLPRGAASVCTGTDMKRLPASPTHLDMLRHLVYGCQVVOGNTL 60
 QY 61 ELTYLPNNAISLSPDIQEVGYTLIAHQVQVPLQRLRVSTQTFEDNYALAVNDNG 120
 DB 61 ELTYLPNNAISLSPDIQEVGYTLIAHQVQVPLQRLRVSTQTFEDNYALAVNDNG 120
 QY 121 DPLNNTTPYTGASGGLRELOLRSLTEILKGVLIQNRPOLCYODTILMKDIFRKNOLA 180
 DB 121 DPLNNTTPYTGASGGLRELOLRSLTEILKGVLIQNRPOLCYODTILMKDIFRKNOLA 180
 QY 181 LTIIDTNRBACHPCSPMKSGSRGSSSDQSLTRTVACGACRKGPLPTDCHEQC 240
 DB 181 LTIIDTNRBACHPCSPMKSGSRGSSSDQSLTRTVACGACRKGPLPTDCHEQC 240
 QY 241 AAGCTGKSHDCLACHFNHSGTEILHCPALVYNTDTFESMPREGRYTFGASCVTACP 300
 DB 241 AAGCTGKSHDCLACHFNHSGTEILHCPALVYNTDTFESMPREGRYTFGASCVTACP 300
 QY 301 YNVLSTVGSCTVLCPLHNOEVTAEQTCRCKSKPCARCYGLGMEHLREVAAYTSAN 360
 DB 301 YNVLSTVGSCTVLCPLHNOEVTAEQTCRCKSKPCARCYGLGMEHLREVAAYTSAN 360
 QY 361 IOEPAGCKKIFGSLAFLPESFDGPASNTAFLQEPOLQVFETLEITGYLISAMPDLP 420
 DB 361 IOEPAGCKKIFGSLAFLPESFDGPASNTAFLQEPOLQVFETLEITGYLISAMPDLP 420
 QY 421 DLSYFQVLYRKGILLNNGAYSLTLOGLISWLGIRSLREISGLALHNHTHLCFVHTV 480
 DB 421 DLSYFQVLYRKGILLNNGAYSLTLOGLISWLGIRSLREISGLALHNHTHLCFVHTV 480
 QY 481 PMDLPFNPHQALHTNRPDECCVGEGLACHOLCARGCHGPGPTCVNCSQFLRGEC 540
 DB 481 PMDLPFNPHQALHTNRPDECCVGEGLACHOLCARGCHGPGPTCVNCSQFLRGEC 540
 QY 541 VEEGRVLYGLPREVYNAHCLPCHPECCOPONGSVTCFPEADQCVACAHYDPPFCVARC 600
 DB 541 VEEGRVLYGLPREVYNAHCLPCHPECCOPONGSVTCFPEADQCVACAHYDPPFCVARC 600
 QY 601 PGVYKPDLSYMPYKPPDEBGAQCPNCTHSCVDLDDKCPABORASPLTSTISAVG 660
 DB 601 PGVYKPDLSYMPYKPPDEBGAQCPNCTHSCVDLDDKCPABORASPLTSTISAVG 660

DB 601 PGVYKPDLSYMPYKPPDEBGAQCPNCTHSCVDLDDKCPABORASPLTSTISAVG 660
 QY 661 ILVVVLGVVFGILLIKRQOKIRKYTWRRLLQETELVEPLTPSGAMENQOMRILKETEL 720
 DB 661 ILVVVLGVVFGILLIKRQOKIRKYTWRRLLQETELVEPLTPSGAMENQOMRILKETEL 720
 QY 721 RAYKVLGSGAFVYVYKGIWTPDENANKIPAIKVLLENTPKANKETLDEAYVWAGVSP 780
 DB 721 RAYKVLGSGAFVYVYKGIWTPDENANKIPAIKVLLENTPKANKETLDEAYVWAGVSP 780
 QY 781 YVERLLGICLSTVQVLTQMPYGLLDHYENRGRGLSDLLNMCQIAKMSYLEDVR 840
 DB 781 YVERLLGICLSTVQVLTQMPYGLLDHYENRGRGLSDLLNMCQIAKMSYLEDVR 840
 QY 841 LVYRDLAARNVLYKSPNHVKITDQGLARLLDIDETEHADGKVPIMWALESTLRRT 900
 DB 841 LVYRDLAARNVLYKSPNHVKITDQGLARLLDIDETEHADGKVPIMWALESTLRRT 900
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEGSELPOPICTIVYMTMKWM 960
 DB 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEGSELPOPICTIVYMTMKWM 960
 QY 961 ISECRPRFRELVSERSRMARDQRFVYIQNEDLGPASPLDSTFYRSLLEDMDGLVDA 1020
 DB 961 ISECRPRFRELVSERSRMARDQRFVYIQNEDLGPASPLDSTFYRSLLEDMDGLVDA 1020
 QY 1021 EETLVPOQGFCCDPAFAGAGVHHRRSSSTRSGGDLTLGLEPSEEARSPPLAPSEG 1080
 DB 1021 EETLVPOQGFCCDPAFAGAGVHHRRSSSTRSGGDLTLGLEPSEEARSPPLAPSEG 1080
 QY 1081 AGSDVFDGDLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVAPLTCSPPQPEYV 1140
 DB 1081 AGSDVFDGDLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVAPLTCSPPQPEYV 1140
 QY 1141 NOPDVRPQPPSPREGPLPARPAGATLERPKTSPGKNGVQVFAFGAVENBEYLTPQ 1200
 DB 1141 NOPDVRPQPPSPREGPLPARPAGATLERPKTSPGKNGVQVFAFGAVENBEYLTPQ 1200
 QY 1201 GGAAPQHPHPAPSPAFDNLVYWDQDEPERGAPSTKGTPTANPEYLGLDVPV 1255
 DB 1201 GGAAPQHPHPAPSPAFDNLVYWDQDEPERGAPSTKGTPTANPEYLGLDVPV 1255

RESULT 2
 TVRNTU
 protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
 C/Accession: A24562; A61204
 R:Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.
 Nature 319, 226-230, 1986
 A/Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
 A/Reference number: A24562; MUID:66118662; PMID:3945311
 A/Accession: A24562
 A/Molecule type: mRNA
 A/Residues: 1-1260 <RAR>
 A/Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746
 R:Meuni, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen, A.; Carlgene 12, 1975-1978, 1991
 A/Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no 1-2-thiazolylformamide or N-methyl-N-nitrosourea.
 A/Reference number: A61204; MUID:92035293; PMID:1682063
 A/Accession: A61204
 A/Molecule type: DNA
 A/Residues: 637-663, 'V', 665-702 <MAS>
 A/Note: authors translated the codon GCA for residue 25 as Val
 C/Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosph
 F/1-19/Domain: signal sequence #status predicted <SIG>
 F/20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

F:658-680/Domain: transmembrane #status predicted <TM>
 F:723-988/Domain: protein kinase homology <KIN>
 F:731-739/Region: protein kinase ATP-binding motif
 F:71,191,263,535,576,634/Binding site: carbohydrate (Aan) (covalent) #status predicted
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted
 F:758/Active site: Lys #status predicted
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 88.0%; Score 5997; DB 1; Length 1260;
 Best Local Similarity 87.9%; Pred. No. 2,5e-237;
 Matches 1105; Conservative 49; Mismatches 101; Indels 2; Gaps 2;

QY 1 MELALCRWGLLALPPGASSTVCTGTDMLRLPASPTHLDMLRLYOGCQVGNL 60
 DB 4 MELAMCRWGLLALPPGASSTVCTGTDMLRLPASPTHLDMLRLYOGCQVGNL 63
 QY 61 ELTYLPTNASLSFLDIQEVGYVLIANQVROPVLPQRLIVRGTOLEFEDNYALAVLDNG 120
 DB 64 ELTYVPANASLSFLDIQEVGYVLIANQVROPVLPQRLIVRGTOLEFEDNYALAVLDNR 123
 QY 121 DPLNNTPTVYTGASPGSLRELQSLTEILKGVLIQNPOLCYODTILMDIFHKNOL 179
 DB 124 DPDQNVAASTPGKRPBGLRELQSLTEILKGVLIQNPOLCYODTILMDIFHKNOL 183
 QY 180 ALTLIDTRSRACHPSCMCKSGRCWGSSEDCSLTRTVGAGGCAACCKPLPTDCCHQ 239
 DB 184 APVDIDTRSRACHPSCMCKSGRCWGSSEDCSLTRTVGAGGCAACCKPLPTDCCHQ 243
 QY 240 CAAGCTGPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPBGRYTFGASCVTAC 299
 DB 244 CAAGCTGPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPBGRYTFGASCVTTC 303
 QY 300 PNYVLTSDVGSCTLVCPFLNNOVTAEDGTORCEKCSKPCARVCYGLGMEHLREVRATSA 359
 DB 304 PNYLTSTVGSCTLVCPFLNNOVTAEDGTORCEKCSKPCARVCYGLGMEHLREVRATSD 363
 QY 360 NIGEPACCKKIFGSLAPLPSPDPDPAANTAPLOPEOLQVETLEITGYLYISAWPDSL 419
 DB 364 NVGEPDCKCKIFGSLAPLPSPDPDPAANTAPLOPEOLQVETLEITGYLYISAWPDSL 423
 QY 420 PDLSTVQNLQVITGRILHANGAYSLTQGLGSIWGLRLSRLBSGLALIHNTILCFVHT 479
 DB 424 RDLSTVQNLRIIGRIIHDAYSLTQGLGSIWGLRLSRLBSGLALIHNTILCFVHT 483
 QY 480 VPMDDOLFRNPHOALLHTANRPDE-CYGBGLACQULCARHCHGPGTQCYNCSQFLRG 538
 DB 484 VPMDDOLFRNPHOALLHTANRPDE-CYGBGLACQULCARHCHGPGTQCYNCSQFLRG 543
 QY 539 ECVEBCRVLQGLPREYVNAHCLPCHEPCOPONGSVTCFPGPADQCAAHYDPPFCVA 598
 DB 544 ECVEBCRVKGLPREYVSDKCLPCHEPCOPONGSVTCFPGPADQCAAHYDSSSCVA 603
 QY 599 RCPSGVKPDLSYPIWKFPDEBGAACOPPCINCHSCVLDLDCGCPAQRASPLTISAV 658
 DB 604 RCPSGVKPDLSYPIWKFPDEBGAACOPPCINCHSCVLDLDCGCPAQRASPLTISAV 663
 QY 659 VGLLVVVLGVVGIILKROQIKRTYWRLLQETELVEPLTPSGAMPNOAMRIKET 718
 DB 664 EGVLLFLVIVVGIILKROQIKRTYWRLLQETELVEPLTPSGAMPNOAMRIKET 723
 QY 719 ELAKVVLGSGAGTGYKGIWIDGENVKIPVAIKYLRNTSPKANEILDEAYVAVAGV 778
 DB 724 ELAKVVLGSGAGTGYKGIWIDGENVKIPVAIKYLRNTSPKANEILDEAYVAVAGV 783
 QY 779 SPVYSLGLCLTSTVQVLTQMLPYGLLDHVENRGRLSGODLLWCMQIAKMSYLED 838
 DB 784 SPVYSLGLCLTSTVQVLTQMLPYGLLDHVENRGRLSGODLLWCMQIAKMSYLED 843
 QY 839 VRLVHRLDLAARVLYKSPHNVKTTDGLARLDIDETRYADGKVPYIKMALESILRR 898
 DB 844 VRLVHRLDLAARVLYKSPHNVKTTDGLARLDIDETRYADGKVPYIKMALESILRR 903
 QY 899 PTHQSDVMSGYVTVWELMTFGAKFPGDIPAREIPDLBNGERLPQPPICITVDYMLMKC 958

DB 904 PTHQSDVMSGYVTVWELMTFGAKFPGDIPAREIPDLBNGERLPQPPICITVDYMLMKC 963
 QY 959 WVIDSECRPRELVESESRNARDPQRFVIONEDLGPASPLDSTFYSLIEDDMGDLY 1018
 DB 964 WVIDSECRPRELVESESRNARDPQRFVIONEDLGPASPLDSTFYSLIEDDMGDLY 1023
 QY 1019 DAEELVLPQGFPCPDPAAGVNHHRNSSLTRSGGDLTLGLEPSEBEAPRSLAPS 1078
 DB 1024 DAEELVLPQGFPCPDPAAGVNHHRNSSLTRSGGDLTLGLEPSEBEAPRSLAPS 1083
 QY 1079 EGAGSDVFDGDLGKAGKGLQSLPTHDSPLQRFSEPTVLPBETDGVVAPLTSRQPE 1138
 DB 1084 EGAGSDVFDGDLGKAGKGLQSLPTHDSPLQRFSEPTVLPBETDGVVAPLTSRQPE 1143
 QY 1139 YVNPQDVRPOPSPBEGPLPAARPAATLERPKTLSPGKGVVDFVAFGAVENPEYLT 1198
 DB 1144 YVNPQDVRPOPSPBEGPLPAARPAATLERPKTLSPGKGVVDFVAFGAVENPEYLT 1203
 QY 1199 PGGAAPOHPHPPAFSAFDNLTYWDDPPERCAPPESTFKCTPAENPEYLGLDV 1255
 DB 1204 PGGTASPPHPPAFSAFDNLTYWDDPPERCAPPESTFKCTPAENPEYLGLDV 1260

RESULT 3
 I48161
 P-185: precursor - golden hamster
 C:Species: Mesocricetus auratus (golden hamster)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
 C:Accession: I48161
 R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa
 Gene 140, 251-255, 1994
 A>Title: Cloning and activation of the Syrian hamster new proto-oncogene.
 A:Reference number: I48161; MIMD:94193007; PMID:7908275
 A:Accession: I48161
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1,1254 <RES>
 A:Cross-references: GB:D16295; NID:9493236; PIDN:BA03801.1; PID:9747595
 C:Genetic:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif

Query Match 87.9%; Score 5993.5; DB 2; Length 1254;
 Best Local Similarity 87.7%; Pred. No. 3.4e-237;
 Matches 1101; Conservative 57; Mismatches 96; Indels 1; Gaps 1;

QY 1 MELALCRWGLLALPPGASSTVCTGTDMLRLPASPTHLDMLRLYOGCQVGNL 60
 DB 1 MELAMCRWGLLALPPGASSTVCTGTDMLRLPASPTHLDMLRLYOGCQVGNL 60
 QY 61 ELTYLPTNASLSFLDIQEVGYVLIANQVROPVLPQRLIVRGTOLEFEDNYALAVLDNG 120
 DB 61 ELTYLPTNASLSFLDIQEVGYVLIANQVROPVLPQRLIVRGTOLEFEDNYALAVLDNR 120
 QY 121 DPLNNTPTVYTGASPGSLRELQSLTEILKGVLIQNPOLCYODTILMDIFHKNOL 180
 DB 121 DPLNNTPTVYTGASPGSLRELQSLTEILKGVLIQNPOLCYODTILMDIFHKNOL 180
 QY 180 ALTLIDTRSRACHPSCMCKSGRCWGSSEDCSLTRTVGAGGCAACCKPLPTDCCHQ 240
 DB 181 ALTLIDTRSRACHPSCMCKSGRCWGSSEDCSLTRTVGAGGCAACCKPLPTDCCHQ 240
 QY 240 CAAGCTGPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPBGRYTFGASCVTAC 300
 DB 241 CAAGCTGPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPBGRYTFGASCVTAC 300
 QY 300 PNYVLTSDVGSCTLVCPFLNNOVTAEDGTORCEKCSKPCARVCYGLGMEHLREVRATSA 360
 DB 301 PNYLTSTVGSCTLVCPFLNNOVTAEDGTORCEKCSKPCARVCYGLGMEHLREVRATSA 360

```

Qy 361 IOEFAGCKKIFGSLAFIPESFDGDPASNTAPLOEQLOVFTLEITGYLYISAMPDLSL 420
Db 361 IOEFAGCKKIFGSLAFIPESFDGDPASNTAPLOEQLOVFTLEITGYLYISAMPDLSL 420
Qy 421 DLSVFNQLOVIRGILNNGAYSLTLQGLISWLGRLSRLSGLSGLALIHNTMLCPVHTV 480
Db 421 DLSVFNQLOVIRGILNNGAYSLTLQGLISWLGRLSRLSGLSGLALIHNTMLCPVHTV 480
Qy 481 PMOQLPNNPQALHTNRPDECEVGBGLACHOCARHCCHGPRPTQCVNCSOFLRQEC 540
Db 481 PMOQLPNNPQALHTNRPDECEVGBGLACHOCARHCCHGPRPTQCVNCSOFLRQEC 540
Qy 541 VEECRVQLGPREYVYARHCLPCHPECOPONGSYTCFEPADQCVACAHYDPFCVARC 600
Db 541 VEECRVQLGPREYVYARHCLPCHPECOPONGSYTCFEPADQCVACAHYDPFCVARC 600
Qy 601 PSQVPLSLYMPIMKPFDEGACOPCPINCTHSCVDLDDKCCPABORASPLTSIIAVVG 660
Db 601 PSQVPLSLYMPIMKPFDEGACOPCPINCTHSCVDLDDKCCPABORASPLTSIIAVVG 660
Qy 661 ILVNVYGVVGLIKRQCKIRKTYMRRLQETELVEPLTPSGAMPQAOMLIKETEL 720
Db 661 ILVNVYGVVGLIKRQCKIRKTYMRRLQETELVEPLTPSGAMPQAOMLIKETEL 720
Qy 721 RKVYVLSGAGVYKGIWIPDGENVKIPVAIKYLRBTSFKANKEILDEAYVYAGVQSP 780
Db 721 RKVYVLSGAGVYKGIWIPDGENVKIPVAIKYLRBTSFKANKEILDEAYVYAGVQSP 780
Qy 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHVENRGRGSLDLMCMQIAKMSYLEDVR 840
Db 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHVENRGRGSLDLMCMQIAKMSYLEDVR 840
Qy 841 LVHRDLAARNLVKSPNNVKTIDPGLARLDDIDETRYHADGKVPKIMALESLIRRF 900
Db 841 LVHRDLAARNLVKSPNNVKTIDPGLARLDDIDETRYHADGKVPKIMALESLIRRF 900
Qy 901 HOSDMSYGVTVVMEIMTFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYMIWCKM 960
Db 901 HOSDMSYGVTVVMEIMTFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYMIWCKM 960
Qy 961 IDSECRPRFRELVEBSFMAADPORFVYIQNEIDLPASPLDSTYRSLLEDMDGLVDA 1020
Db 961 IDSECRPRFRELVEBSFMAADPORFVYIQNEIDLPASPLDSTYRSLLEDMDGLVDA 1020
Qy 1021 BEYLVPOQGFPCPPAPAGAGMHHRRSSSTRSGGGLUTGLBPSSEKARSPLABSEG 1080
Db 1021 BEYLVPOQGFPCPPAPAGAGMHHRRSSSTRSGGGLUTGLBPSSEKARSPLABSEG 1080
Qy 1081 AGSDVPFODDLGMAKGLQSLPTHDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYV 1140
Db 1081 AGSDVPFODDLGMAKGLQSLPTHDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYV 1140
Qy 1141 NQPEVRROPPLTPGCPPLPVPAAGATLERPKTLSPGKGVYKVDYTFGGAIVENPEYLVPR 1200
Db 1141 NQPEVRROPPLTPGCPPLPVPAAGATLERPKTLSPGKGVYKVDYTFGGAIVENPEYLVPR 1200
Qy 1201 GGASQDPH-PPALCPAPDNLTYMDDPSEBRSPPATFPGTFAENPEYLGIDVY 1254
Db 1201 GGASQDPH-PPALCPAPDNLTYMDDPSEBRSPPATFPGTFAENPEYLGIDVY 1254

```

RESULT 4

epidermal growth factor receptor precursor - human
 N:Contains: protein-tyrosine kinase (BC 2.7.1.112) erbB
 C:Species: Homo sapiens (man)
 C:Date: 15-Nov-1994 #sequence revision 27-Nov-1995 #text change 11-Jun-1999
 C:Accession: A00641, A25772, E30024, A38672, A00642, A03655, A23062, A05281, A60143, A33
 R:Ulrich, A.; Condeelis, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Jam, A.W.; Lee, J.; Y
 rg, P.H.
 Nature 309, 418-425, 1984
 A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of

A:Reference number: A00641, MUID:84219729, PMID:6328332
 A:Accession: A00641
 A:Molecule type: mRNA
 A:Residues: 1-1210 <LIN>
 A:Cross-references: EMBL:X00588, NID:g31113, PIND:CAA25240.1, PID:g757924
 A>Note: The authors translated the codon AAG for residue 540 as Asn
 R:Jin, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Mellino, G.T.; Pastan, I.
 Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985
 A:Title: Characterization and sequence of the promoter region of the human epidermal gr
 A:Reference number: A25772, MUID:85270438, PMID:2991899
 A:Accession: A25772
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-29 <ISH>
 A:Cross-references: GB:M11234, NID:g181981, PIND:AAA2370.1, PID:g553272
 R:Hailey, J.; Whittle, N.; Bennett, P.; Kitchington, D.; Ulrich, A.; Waterfield, M.
 Oncogene Res. 1, 375-396, 1987
 A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification
 A:Reference number: S30024, MUID:88217333, PMID:3329736
 A:Accession: S30024
 A:Molecule type: DNA
 A:Residues: 1-29 <HA2>
 A:Cross-references: EMBL:X06370, NID:g31118, PIND:CAA29668.1, PID:g31119
 R:Hailey, J.D.; Waterfield, M.D.
 J. Biol. Chem. 266, 1746-1753, 1991
 A:Title: Contributory effects of de Novo transcription and premature transcript termina
 A:Reference number: A38672, MUID:91107677, PMID:1988448
 A:Accession: A38672
 A:Molecule type: DNA
 A:Residues: 1-29 <HAL>
 A:Cross-references: GB:A38425, NID:g181977, PIND:AAA63171.1, PID:g553271
 A:Experimental source: carcinoma cell line A431-7
 R:Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.; Me
 Nature 309, 806-810, 1984
 A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs
 A:Reference number: A00642, MUID:84245835, PMID:6330563
 A:Accession: A00642
 A:Molecule type: mRNA
 A:Residues: 799-811, 'R', 813-942 <XUY>
 A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF recep
 R:Lin, C.R.; Chen, W.S.; Krutiger, W.; Stolaretzky, L.S.; Weber, W.; Evans, R.W.; Verna, I
 Science 224, 843-848, 1984
 A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplification
 A:Reference number: A43615, MUID:84196372, PMID:63326261
 A:Accession: A43615
 A:Molecule type: mRNA
 A:Residues: 713-964 <LIN>
 A:Experimental source: epidermoid carcinoma cell line A431
 R:Stamen, F.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.
 Biochem. Biophys. Res. Commun. 124, 125-132, 1984
 A:Reference number: A23062, MUID:85046483, PMID:6093780
 A:Accession: A23062
 A:Molecule type: mRNA
 A:Residues: 1028-1210 <SIM>
 R:Weber, M.; Gull, G.N.; Speiser, J.
 Science 224, 294-297, 1984
 A:Reference number: A05281, MUID:84172183, PMID:6324343
 A:Accession: A05281
 A:Molecule type: protein
 A:Residues: 25-30, 'S', 32-51, 454-467 <WEB>
 R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.
 J. Biol. Chem. 260, 5205-5208, 1985
 A:Title: Identification of residues in the nucleotide binding site of the epidermal gro
 A:Reference number: A60143, MUID:85182650, PMID:2985580
 A:Accession: A60143
 A:Molecule type: protein
 A:Residues: 740-744, 'X', 746-747 <RUS>
 R:Motczkowski, B.; Mosig, G.; Cohen, S.
 Nature 309, 270-273, 1984
 A:Title: ATP-stimulated interaction between epidermal growth factor receptor and superc
 A:Reference number: A38023, MUID:84191554, PMID:6325948
 A:Contents: annotation; receptor activity

A/Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA
 R/Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.
 Cell 59, 33-43, 1989
 A/Title: Functional independence of the epidermal growth factor receptor from a domain
 A/Reference number: A3331; MIMD:9003233; PMID:2790960
 A/Content: annotation, internalization signal
 C/Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor
 C/Genetics:
 A/Gene: GDB:EGFR
 A/Cross-references: GDB:120610; OMIM:131550
 A/Map position: 7p12.3-7p12.1
 C/Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP, autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
 P/1-24/Domain: signal sequence #status predicted <SIG>
 P/25-1210/Product: EGF receptor #status predicted <MAT>
 P/25-645/Domain: extracellular #status predicted <EXT>
 P/75-300/Domain: EGF receptor extracellular domain repeat <E2>
 P/390-600/Domain: EGF receptor extracellular domain repeat <E2>
 P/646-668/Domain: transmembrane #status predicted <TM>
 P/669-1210/Domain: intracellular #status predicted <INT>
 P/710-975/Domain: protein kinase homology <KIN>
 P/718-726/Region: protein kinase ATP-binding motif
 P/999-1046/Region: coiled-coil mediated internalization signal
 P/1047-1210/Region: inhibitory
 P/128,175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status predic
 P/745/Active site: Lys #status experimental

Query Match 46.54; Score 3169; DB 1; Length 1210;
 Best Local Similarity 49.84; Pred. No. 4e-122; Indels 116; Gaps 22;
 Matches 632; Conservative 177; Mismatches 345;

```

11 LLLALLPPGAA--STQVCTGDMKRLPASPELIDMLRLHYOCQVVOGNTLTYLPTN 68
14 LLAALCASALBEKKYCCQGSNTLTLQGFEDHFLSLQRMNNCEVLLGLLITTYQBN 73
69 ASISFLDIOEVQGYLIANQVQLQRLRYRGVLQPEDNYALVLDNGPLNNTTP 128
74 YDSFLTIQVAVGYLIANTVERIPLENLIIRGMYYENSYALVLSYD----- 126
129 VTQASPGALRELRSLTEILKGGVLLQBNPOLCYODTILMKDIFHNNQALTLIDTNR 188
127 ---ANKTGLKELPRNLQELHGVNRPSPNPNALCNVSIQWRDVSSDPLSNMSMDPNH 183
189 SRACHPCSPKCSGRSCWSESSDQSLTRTYCAGCA-RCKGPLPTDCHCQCAAGCTG 247
184 LSCCQKQDPCSPKSCWAGBENQKTKITICAGCCGRCKGKSPSCINQCAAGCTG 243
248 KHSDDLCLAHFNHSGICELCPALVTYNTDTFESMPNREGRTFGASCVTACPNYLSTD 307
244 RESDCLVCRKFRDEATCKDTCPLMLYNPTTYQMDVNPBGKYSFGATCVKKCPNRYVTD 303
308 VGSCITLVPLHNOEVTADGTQREKSKPCARVCYGLGMEHLREVAATVANSIQERAGC 367
304 HGSCTVACGADSYEM-EDGVKCKCKGPRKVCNIGIGIFXDSLSINATNKHFKNC 362
368 KCKFGSLAPLPESGDPASNTAPLOPRLQVFTLEITSGYLYTISAMPDPLSLYFON 427
363 TSISGDLHILPVARGDSFTHTPLDPOEDILTKVKEITSTFLLIQAMPENRDLHAFEN 422
428 LQVIRGRILHNGAVSLTLQGLISWLGRLSLRELSGSLALIHNTHLCFVHTVPMDFR 487
423 LEIIRGRTKHGGSLAVSLNITSIGRSIKELISDGVIIISGNKNCYANTIMKGLFG 482
488 NPHQALHTANRPDECVGSLAHOICARHCGPFGTQVCVNSQPLRGCECYBEGRVL 547
483 TSGQKTKIINRGNKSCATQVCHALCPGCGPFRDVSGRNYSRGEBCVCKKTL 542
548 OGLPREVYNARHCLPCECOPONGSYTCFGBADOCACAYADPPCVARCPSPGYPD 607
543 EGBREPREVNSCTQCHPECLPOMNITTCRGPDNQCCHNYIDGPHCVTCAGAWGE 602
608 LSTVPIWKPDEBAGCPCPINCHSGVDLDDKSCPAQRASPLTISAVG---ILLV 664

```

```

DB 603 NNTL-VWKYADAGVCHLCHPNTCTGCTGPGLSCTPNPKIP--STATGVGALLLLV 659
QY 665 VLVGVVFGILIKRQOKIRKYTWRRLLQETELVEPLTPSGAMPNOQWRILKETELRYK 724
DB 660 VALGIG---LPMRRRHVRKTRLRLLQRELVLEPLTPSGAPNOALLRIKETEFKKIX 716
QY 725 VLSSGAFGYKGIWIPDGENVKIPVAIKYLRNTSPKANKELIDEAVVMAGVSPVYSR 784
DB 717 VLSSGAFGYKGIWIPDGENVKIPVAIKYLRNTSPKANKELIDEAVVMAGVSPVYSR 776
QY 785 LIGICLSTVQVLTQMPVCCCLDHYENRGRGSDOLMMQOIAQMSLDEVDRLVHR 844
DB 777 LIGICLSTVQVLTQMPVCCCLDHYENRGRGSDOLMMQOIAQMSLDEVDRLVHR 836
QY 845 DLAAKRVLYSPNHNKTTDGLARLLIDETEVHADGAKYPIKMMALLESILRRFTHQSD 904
DB 837 DLAAKRVLYSPNHNKTTDGLARLLIDETEVHADGAKYPIKMMALLESILRRFTHQSD 896
QY 905 VMSGYTVMELMTFGAKPYDGIPIAREIPDLLEKGERLPPICTIDVYIMVKKMIDSE 964
DB 897 VMSGYTVMELMTFGAKPYDGIPIAREIPDLLEKGERLPPICTIDVYIMVKKMIDSE 956
QY 965 CRPRRELVESEPRMADPQRFVITQ-NEDLPASPLDSTFYRSLDEDDMDGLVDAREY 1023
DB 957 SRPRFRLIEFSRMADPQRFVITQ-NEDLPASPLDSTFYRSLDEDDMDGLVDAREY 1016
QY 1024 LVPQGFECFCDPAPAGAGMHHRRSSSTRSGCDLTLGLEPSEEAAPRSPLASEGAGS 1083
DB 1017 LVPQGFECFCDPAPAGAGMHHRRSSSTRSGCDLTLGLEPSEEAAPRSPLASEGAGS 1042
QY 1084 DVPDGLGMAAGLQSLPETHDSPLQRYSEDPVLPSET--DGYVAALPSCPOPEYVN 1141
DB 1043 N--NSTVACIDRNGLOSCPIKEDSFLQRYSSDPGALTSEIDDTFL-----PVPEYIN 1094
QY 1142 QPDVRRQPPSPRGRPLPAAPACATL-----ERPKTSLPGNGGVVQVAFAGAVENREY 1196
DB 1095 Q-----SVF-KRPASGVQNPVYHNOPLNAPSPSDPHYD--PHSTAGNPEY 1138
QY 1197 L-TPQGAAPQHPHPFASPAFDNLYWDO-----DP-----PERGAPSTPEK 1239
DB 1139 LNTVQ-----PFCVNSTFDSPAHMAQKSHQISLDNPDYODPFPRKAKRNGIRKG 1189
QY 1240 TPTAENPEYL 1249
DB 1190 S-TAENAEYL 1198

```

RESULT 5
 A53183
 A/Title: epidermal growth factor receptor precursor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 18-Jun-1999
 C/Accession: A53183; A43818; S24942; A28941; S45325; 149643
 R/Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Eard, H.S.; Jenkins, N.A.;
 Genes Dev. 8, 399-413, 1994
 A/Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor
 A/Reference number: A53183; MIMD:94110986; PMID:8125255
 A/Accession: A53183
 A/Molecule type: mRNA
 A/Residues: 1-1210 <LUB>
 A/Cross-references: GB:U03425
 R/Avivi, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.
 Oncogene 6, 673-676, 1991
 A/Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site
 A/Reference number: A43818; MIMD:91232866; PMID:2030916
 A/Molecule type: mRNA
 A/Residues: 1-714 <AVI>
 A/Cross-references: GB:X59698
 R/Bislinger, D.P.; Serrero, G.
 submitted to the EMBL Data Library, June 1992
 A/Reference number: S24942
 A/Accession: S24942

A/Molecule type: mRNA
 A/Residues: 969-971, 'K', 973-1115, 'D' <EIS>
 A/Cross-references: EMBL:Z12608
 R/Heiser, G.J.; Gili, G.N.
 J. Biol. Chem. 263, 13152-13158, 1988
 A/Title: Epidermal growth factor receptor threonine and serine residues phosphorylated
 A/Reference number: A28941; MUID:80330814; PMID:3138233
 A/Accession: A28941
 A/Molecule type: protein
 A/Residues: 689-694, 'X', 696-704, 'L', 706-707, 989-992, 'XX', 995-996, 'X', 998-1000, 1002-1009, submitted to the EMBL Data Library, April 1994
 A/Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor
 A/Reference number: S45325
 A/Accession: S45325
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1971, 'K', 973-1210 <VER>
 A/Cross-references: EMBL:X78987; NID:9488830; PIDN:CA55587.1; PID:9488831
 R/Parla, B.C.; Das, S.K.; Andrews, G.K.; Day, S.K.
 Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
 A/Title: Expression of the epidermal growth factor receptor gene is regulated in mouse
 A/Reference number: I49643; MUID:93126380; PMID:7678348
 A/Accession: I49643
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 12-20, 22-132 <RES>
 A/Cross-references: GB:I06864; NID:9193001; PIDN:AAA53029.1; PID:9567201
 C/Genetics:
 A/Genes: EGR1
 C/Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP; growth factor receptor; kinase-related transforming protein; phosphoprotein
 F/1-24/Domain: signal sequence #status predicted <SIG>
 F/648-670/Domain: transmembrane #status predicted <TM>
 F/712-977/Domain: protein kinase homology <KIN>
 F/720-728/Region: protein kinase ATP-binding motif
 F/680, 695/Binding site: phosphate (Thr) (covalent) #status experimental
 F/697, 1070, 1071/Binding site: phosphate (Ser) (covalent) #status experimental
 F/993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
 F/1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental
 F/1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 46.1%; Score 3145; DB 2; Length 1210;
 Best Local Similarity 49.7%; Pred. No. 3,86-121;
 Matches 635; Conservative 169; Mismatches 353; Indels 120; Gaps 24;

11 LLLALLPPGA--STOYCTGTMDLRLPASPETHLMDRLHYGCGOVQGLLETLPTN 68
 14 LTRLCAGALREKVKCOGTSNRLTOLGTYEDHPLSLQRMYNCEVVLGLLETTYORN 73
 69 ASLSFLDIOEVQGYVLIAHQVROVPLQRLIVRGTOLEPDNYALAVLDNGDPLANTTP 128
 74 YDSFLKTEIGEVAVYLIANTVERIPLENQIIRGNALYENTYALAILSN----- 124
 129 VTGASPGGLREBLQRLSTLEIKGVLIQRPOLCYODTILMKDI---FKKNQALATLI 184
 125 -YGTNRGTGLRELPRKNIQELIIGAVRFSNNPILCMQDITQWRDIVOVNFMSNMMDL-- 180
 185 DYNRSRACHPSCPMCKSGKSGESSEDCSLTRTYVCAAGCA-RCKGPLPTDCCHGCGAAG 243
 181 -QSPSSCPKCDPSCPNCSGCGGSENCQKLTIKIACQCSHRCKSGSPSCCHNQCAAG 239
 244 CTGPKHSDCLACLPFNHSGICELCPALVTYNTDTFBSMPNREGRTFGASCVTACPPNY 303
 240 CTGRREDDCLVQCFQDEBATKDCPRLMLNPTTYQDVNPKGYSGATCVKCKPPNY 299
 304 LSTVSGCTLVCPILNQEVTLEDGTQCKCKSKPCARVYGLGHEHLREAVTASNIQ 363
 300 VVTHGSCVRAKCGDYDEV-EDGIRKCKCKDGCRCRKNIGIGIFGFTLSTINATNIK 358
 364 FAGKKIKFGLSLAPLPSFPDGPASNTAPLOEOLQVFTLEIRGYLYISAMPSTPLDS 423
 359 FKYCTAISGDLHILPVAKGDSPTKTPPLDRELLEILKTVREITGFLTIQAMPDNMTDL 418

424 VFQNLQVIRGRIILANGAVSLTLOGLISWGLSLRELSSGLALIHNTHLCEVHTVPM 483
 419 AFENLEIRGRTRKHQOFSLAVVGLNITSLKSLKEISDGDVILSGNRILCVANTINWK 478
 484 QLRNPHQALLHNRNPEDECVBGLAQHQLCARHGWGSPPOCVNCSQFLRQECVEE 543
 479 KLFETPQKTKIMNNNAEKCKAVNHCNPLCSBGCMGEPDVCSCVNSRRECEVEK 538
 544 CEVLQGLPREYVARRHCLPCHPECOPONGSFTCFEPADQCVACAHYKDPFCVACPSG 603
 539 CHLLEGPFRFVENSECQHPCELPQAMNITTCGRPDNICCAHITDEHCYKTCBPAG 598
 604 VKEDLSYMPIKRPDEBACQPCPINCTHSCVDLDKXCPABQASPLYSISAVGILL 663
 599 IMENNTL-VMKYADANNVCHLCHACTYGCAGRGLOGCEWPSGPKIPISATGIVGILL 657
 664 VVVLGVVFGI-LIKRQOKIRKTYTMRLLQETLYVEPLTSSGAMPNOAMILKETEIRK 722
 658 FIVV-VALGIGLPMRRRHVIRKKTLLRLQERLEVEPLTSSGEPNQAHLRIKETEIRK 716
 723 VKVLGSAFGTYVKKGIWIPDGENVKIPVALKVLRENTSPKANKELDEAVYMAVGSPYV 782
 717 IRLVSGARFTYVKKGLMIPGEEKYKIPVALKELREASIPKANKELDEAVYMAVDNHN 776
 783 SRLIGICTSTVOLVTOIMBYGCLLDHVRNRRGLSSQDILLNMCQIAGKMSYLEDVRLV 842
 777 CRLIGICTSTVOLVTOIMBYGCLLDVVRHKNIGSQYLLNMCQIAGKMSYLEDVRLV 836
 843 HEDLAARNVYKSPNHYKITDPRGLAATLDLDEHYHADGKVPKXMALESILRRFRHQ 902
 837 HEDLAARNVYKSPNHYKITDPRGLAATLDLDEHYHADGKVPKXMALESILRRFRHQ 896
 903 SDVWSYGVYTWELMTFGAKPYDGIPIAREIDLEKGRLLPOPICTIDVYIMYKCMID 962
 897 SDVWSYGVYTWELMTFGAKPYDGIPIAREIDLEKGRLLPOPICTIDVYIMYKCMID 956
 963 SECPREPRELVSEFSRMAKDPORFVITQ-NEDLGPASPLDSTFYSLLEDQMDGLVDAE 1021
 957 ADRBRKRELITLFSQARDPQRYLVIQGDERMILPPTDSNFPYALMDEEDMDVDVAD 1016
 1022 EYLVPOGFRCPDPAFAGGVNHRHSSSTRSGGDLTGLSESEEARPSPLAPSGA 1081
 1017 EYLVPOGFRCPDPAFAGGVNHRHSSSTRSGGDLTGLSESEEARPSPLAPSGA 1042
 1082 GSDVFDGDLGMAKAGLSLPTDPSPLQRYSEDPVPLPSET---DGVAPLTPCSPOPEY 1139
 1043 TSN---NSTVACINRNGSCRYKEDATLQRYSSDPFGAVTBDNIDDAFL-----PVREY 1092
 1140 VNQPDVRRPSPRPGPLPAPAPAGATL---BRPTLSFGKNGVVDVAFGAVENP 1194
 1093 VNQ-----SVP-KRPAQSVQNPVYHQPLHAPAGRLHYQN--PHSAVAVNP 1136
 1195 EYL-TPOGGAAPQHPRPAPSPAFDNLYWDQ-----DP-----PERGAPRPTF 1237
 1137 EYLVTAQ-----PTCLSSGFSNPAALMTQKSHQMSLDNPYQODFPPEKTPKNGIF 1187
 1238 KGPPTAENPEYLGIDVP 1254
 1188 KG-PTAENAEYLVAVP 1203

RESULT 6
 T/CHLV
 epidermal growth factor receptor precursor - chicken
 N/Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB
 C/Species: Gallus gallus (chicken)
 C/Date: 28-Feb-1986 #sequence revision 05-May-1995 #text change 04-Feb-2000
 C/Accession: A27720; A00643
 R/Lax, I.; Johnson, A.; How, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Vennet;
 Mol. Cell. Biol. 8, 1970-1978, 1988
 A/Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in mo
 A/Reference number: A27720; MUID:88261272; PMID:3260329

[illegible]

Db 542 CVKQCNILQSGSRREFRFBDSKCLPCHSBLCVONSTAYNTTCSGSPGDHCKMKAHFLDGRPC 601
QY 597 VARPSPGVKPDLSMTPLTKPFPDEBGACQPCPINCSTSCVDLDDKCCPABORASPLTSTIS 656
Db 602 VKACPAGVGLGENDTL-VWKYADANAVCOLCHPNCTRGCGKGLBEGCP--NGSKTPPSIA 657
QY 657 AMV-GILLVVLGVVPGILIRBROOKIKRYTMRRLTOETLVEPLTPSGAMPNOAMRIL 715
Db 658 GVAGGLCLLVVVGIGIGLYLRR-HIVAKRTLRRLDRELVEPLTPSGEABNOAHLRL 716
QY 716 KETELRKVKVLGSGAFGVYKGIWIPDENYKIPVAIKVLENTSPKANKETLDEAYVMA 775
Db 717 KETEFKVKVKVLGSGAFGVYKGIWIPGEKVKIPVAIKVLENTSPKANKETLDEAYVMA 776
QY 776 GVSGPYVSRLLIGLITLSTVQLVTLQMLPFGCLLDHVRKRGSLGSDLLNMCMQIAKGMXY 835
Db 777 SVNDPHVCRLLIGLITLSTVQLITQMLPFGCLLDYIRHKDNIISGYLLNMCMQIAKGMXY 836
QY 836 LEDVRLVHRDLAANVAVKSPNHYKITDPGLARLLDIDETRYHADGGKVPKIMALESTL 895
Db 837 LEBERLVHRDLAANVAVKTPQHKITDPGLAKLLGADBEKTYHABGKVPKIMALESTL 896
QY 896 RRRFTHSDVMSYGVYVWELMTFGAKPYDGIIPAREIPDLLEKGERLPORPCTIDVYIM 955
Db 897 HRIYTHOSDVMSYGVYVWELMTFGSKPYDGIIPASEISSVLEKGERLPORPCTIDVYIM 956
QY 956 VKCMIMDSKPRRELVSERSBARDPQRVVVQ-NDLDGPASLDSTFYSRLLEDOM 1014
Db 957 VKCMIMADSPPKRELIAEFSKMAKDPRLVLYQGDRIHMLPSTLDSKPRFTLWEEEM 1016
QY 1015 GDLVDAEYVLPOOGFFCPDPAPGAGGVHHRHSSSTRSGGDLTLGLPESEEPARSP 1074
Db 1017 EDYDADAEYVLPHGFF-----NSPST-----SRTP 1042
QY 1075 L-----APSEAGSDVDPDGDIMGAKAGLQSLPHNDPSPLORYSEDPVLPSEF--DCY 1127
Db 1043 LLSLSLATSNNSAATNCID-----RNGQGHPRVEDSFVORYSDPFGNLEESIDDF 1094
QY 1128 VAPLITCPOPEYVNOPIVRPQPPSPRSGRLPAARPAAGTLERPKTLSSGKXGVVDVF-- 1185
Db 1095 L-----PAPRYVNO-LMPKPS-----TAVVQVQIYNNISLT 1125
QY 1186 -----AFGGAENPEYLTPOGGAAPORPHPRPAESPAPDNLVYMDQ----- 1225
Db 1126 ALSTLPMDSRKQNSHSTRVNDPEYL-----NTNGSPFLAKTYPESSPYMQSGNHQIN 1177
QY 1226 -DPPE-----RGAPSTFGKTPTAENPEYLGDPV 1254
Db 1178 LDNDPYOODFLPNETKPNGLLKVPALENPEYLRVAAP 1214

RESULT 7
A47253
epidermal growth factor receptor, HER4 - human
C|Species: Homo sapiens (man)
C|Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C|Accession: A47253
R|P|L|Man, G.D.; Colousecu, J.M.; Whitney, G.S.; Green, J.M.; Carlton, G.W.; Foy, L.; N
Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993
A|Title: Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal
A|Reference number: A47253; MUID:93189574; PMID:8383326
A|Accession: A47253
A|status: preliminary; not compared with conceptual translation
A|molecule type: nucleic acid
A|Residues: 1-1308 <EEO>
A|Cross-references: GB:107868; NID:9337359; PIDN:AA55946.1; PID:9337360
A|Note: sequence extracted from NCBI backbone (NCBIP:126842)
C|Superfamily: epidermal growth factor receptor; protein kinase homology
C|Keywords: ATP; growth factor receptor
P:716-981/Domain: protein kinase homology <KIN>
P:724-732/Region: protein kinase ATP-binding motif

Query Match 44.1%; Score 3004.5; DB 2; Length 1308;
 Best Local Similarity 45.6%; Pred. No. 2,26-115;
 Matches 614; Conservative 183; Mismatches 375; Indels 175; Gaps 27;

9 MGILLALLPPGA-----STVCTGTDMKRLPASPEHMLRLHLYGSCOVVGNLTETLY 64
 8 WWWSLVLAAGTVPSPSOSVAGTENTLSLDBEQYALRYTENCERVWGNLTITS 67
 QY 65 LPTNASLFLQDIOEVGYVLIANOVROVPLQRLIRVGTQLEFEDNYALAVLDNGPLN 124
 Db 68 IEHRDLSFLRSVRELVGYVALNQPFLYLENLRIRIGTKLVEDRALALIFLNYRKG 127
 QY 125 NTTPVTGASREGLEQLRLSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNQALTLI 184
 Db 128 NF-----GLOELGKLTLEILNGGVYDQNKFLCYADTIIMODIVRNPMPSNLTIV 178
 QY 185 DTRSRACHPSPMCKSRCKGSESDCOSLITRVACGGC-ARCKGPLPDDCCEQAG 243
 Db 179 STNGSSCGGRHKSCTG-RCWGPTEHNCQTLIRTVCAEQCGRCYGVSDCCRECGAG 237
 QY 244 CTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNREGRYTFGASCYTAPEYNY 303
 Db 238 CSGKPDIDCRACNNPNDGACVTCQPTQPVYNNPTTPOLEHNPNAKITYTGACVCKCPHNF 297
 QY 304 LSTDVSGCTLVCPHNOEVTAEDETQRCCKSKPCARVCGLGMEHLREVAVTSANIOE 363
 Db 298 V-VDSSSCVACPSKMEV--EENGIKMKCPCTDICKPKACDGICTGSLMSAQTVSSNDK 355
 QY 364 PACGKTRFGLAFLPESFPDGPANTAPLOEQIQLPTELEITGYLISAMPDSDPLS 423
 Db 356 FINTKTNGLTFLVTGIGHDPYALIEAIDCKLVNFTVEITGFLMIQWSPMPMDFPS 415
 QY 424 VFORLQVIRGRILHNGAVSLTLOGLISWLGSLRLSRLSGSLALIHNTHLCPYATVPMD 483
 Db 416 VFSVLVIGRVLVYSGSLTLILKQGGITSLOPGLKEISAQNIYITNSNLCYHTINMT 475
 QY 484 QLPNPQALHTANRPDECEVGGGLACHQLCANGHCKPPTQCVNCSQPLRGQECYEE 543
 Db 476 TLFSTINQRIYIRNRKAENCTAGAMVCNHLCSGSGCGPDPDCLSCRRFSRGICIES 535
 QY 544 CRVLQGLPREVYVNRHCLPCHRECPOR-ONGSVTCGPRADQVCAHAKDPPCYARCP 602
 Db 536 CNLTGDEREPENSGICVECDPCKEKGEDGLTCHGEPDPDCTKSHKDPGNCCEKPD 595
 QY 603 GVXKDLSTYPIWKPEDEGACOPRPINCTHSCVLDLDKGC-----PABORASPL 651
 Db 596 GLOGANF--IFKXADPRBCHPCHPNTQGCNPTSHDCIYYPWTGHSITLPOHAR----- 649
 QY 652 TSITISAVV--GILLVVLGVVFGILIKRQOKIRKTYTRILLQETELVEPLTPSGAMPQ 709
 Db 650 TPLIAGVTGGLFTLIVIVGLTFYAVVRKSIK-KKRALRRFL-ETLEVEPLTPSGTAPNQ 707
 QY 710 AOMELKTELRKTKVYGGAGFVYVYKIMV PDGNNVYKIPALVYLRNTPSPKXKELTD 769
 Db 708 AQLILKETEKRKYVSSGAFVYVYKIMV PEGTAVKIPALVYLRNTPSPKXKELTD 767
 QY 770 EAYVAGVSPYVRLIGITCLSTVQVLTQMLPYGCLLDHYRENGRGLSODLLNMCQI 829
 Db 768 EALIMASMDHVLVRLGVCSPITQVLTQMLPHGCLLEVYHNHNDNGSOLLNMCQI 827
 QY 830 AKGMSYLEDVRLVHRDLAARVLYKSPNHVKITDPGLARLLDIDETEHADGKVPYKIM 889
 Db 828 AKGMSYLEDVRLVHRDLAARVLYKSPNHVKITDPGLARLLDIDETEHADGKVPYKIM 887
 QY 890 ALBSILRRFTHOSDWSYGVTVWELMFGAKPYDGTIPAREIPOLLKGBRLPORPCTI 949
 Db 888 ALBSILRRFTHOSDWSYGVTVWELMFGAKPYDGTIPAREIPOLLKGBRLPORPCTI 947
 QY 950 DVYVIMYKCMNIDSECRPRELVESEFARMADPORFVYIQNED-LGSPPLDSTFYRSL 1008
 Db 948 DVYVIMYKCMNIDSECRPRELVESEFARMADPORFVYIQNED-LGSPPLDSTFYRSL 1007
 QY 1009 LEDDMQDLVAEETLVYQGGPCPPDPAFGAGVYVHNRHSSSTRSGGGDITLGLPSEEE 1068

Db 1008 LDEEDLEDMWDAEETLVYV-QAFNIPP-----LYTSRARDISNRS-----EIGSPPPA 1055
 QY 1069 EAPRS-----PLAP-SEGAGSDVFPDGLGMGAALQ 1099
 Db 1056 YTPMSGNQPYVRDGFAPAEQGVSVYPRAPSTITIEAVVAQATRAIFPDSCNGTLRKPFV 1115
 QY 1100 SLPTHDSPLQRYSEDPFTVPLPS-----ETDGYVAPLTCSPQPEYVNOQDVNRQPQSP 1152
 Db 1116 APHQEDSSQRYSADETFVAPERSPRGLEDSEGYMTPMRKPKQEXYLNPYE----- 1167
 QY 1153 REGPLAARAGATLEHPTKLSRCKNGVYDVAFGCAVENEPEYLTPOCGAARPOHPPA 1212
 Db 1168 -ENPFVSR-----KNGDLO-----ALDPEYHNASNG-----PPKA 1198
 QY 1213 -----FSPAFDLYVWDDPPERGA--PPST 1236
 Db 1199 EDEYNEPPLYANTPANTLGAKEYLKNILSMPEKAKKAFDNPDPYNNHSLPFRSTLQHDPY 1258
 QY 1237 FKGTPT-----AENPEYL 1249
 Db 1259 LQESTKCYFKONGRIRPIVAENPEYL 1285

RESULT 8
 506142
 protein-tyrosine kinase (BC 2.7.1.112) mrk-y precursor - southern platyfish
 N:Alternate names: epidermal growth factor receptor homolog; kinase-related transformin
 C:Species: Xiphophorus maculatus (southern platyfish)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 18-Feb-2000
 C:Accession: S06142; S15809
 R:Witdbprod, J.; Adam, D.; Maltischeck, B.; Maenele, W.; Raulf, F.; Telling, A.; Robert
 Nature 341, 415-421, 1989
 A:Title: Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu 10
 A:Reference number: S06142; MUID:90015140; PMID:2797166
 A:Accession: S06142
 A:Molecule type: DNA
 A:Residues: 1-1166 <MIT>
 A:Cross-references: EMBL:X16891; NID:965290; PIDN:CAA34770.1; PID:965291
 R:Adam, D.; Maenele, W.; Scharf, M.
 Oncogene 6, 73-80, 1991
 A:Title: Transcriptional activation of the melanoma inducing Xmrk oncogene in Xiphophor
 A:Reference number: S15807; MUID:91125882; PMID:1846957
 A:Accession: S15809
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 821-1025, 'N', 1027-1098, 'A', 1100-1166 <ADA>
 A:Cross-references: EMBL:X56319; NID:965284; PIDN:CAA39763.1; PID:965285
 C:Genetics:
 A:Gene: mrk
 A:Map position: Y
 A:Intons: 872/3; 898/1; 947/1; 979/3; 1025/3; 1056/1
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor; phosphotransferase; transmembrane protein; tyr
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:126-1166/Product: kinase-related transforming protein (Tu) #status predicted <MAT>
 F:707-972/Domain: protein kinase homology <KIN>
 F:715-723/Region: protein kinase ATP-binding motif

Query Match 39.7%; Score 2708; DB 1; Length 1166;
 Best Local Similarity 45.5%; Pred. No. 2,36-103;
 Matches 576; Conservative 164; Mismatches 389; Indels 138; Gaps 28;

4 AALCRMGILLALLPPGAAT-----QVCTGTDMKRLPASPEHMLRLHLYGSCOVVGN 59
 Db 8 AALQO--LLVLVLSRCCSTDPDRKQCGTSNQMTM---LDNHVYKMKKMTSGCNVLEN 62
 QY 60 LEITLPTNASLFLQDIOEVGYVLIANOVROVPLQRLIRVGTQLEFEDNYALAVLDN 119
 Db 63 LEITYOENDLSFLQSIQEVGYVLIAMNEVSTIPLVNLRILGONLYEGNFTLLVMSN 122
 QY 120 GDPINNTTPVTGASPGGLRELQLASLTREILKGGVLIQRNPOLCYODTIIMKDI FHKNNQ 179


```

Db 123 YOK-NPSSP--DVYQVGLKQLOLSNLTLLSGGVKSHNPCLCNVETIMMDIYDXTSNP 179
Qy 180 ALTLIDYNSRAHCPSPCKSGSRGWESSBDCQSLTRTVCAAGC--ARCKPLPTDCHE 238
Db 180 TNNLTPHAEBCQKCDHCNVAGSCWAPRGHCQKTKLLCAEQCRRCRGRPRIDCCHE 239
Qy 239 OCAAGCTGKXSDCLACLHPNHSIGELHCPALVTYNTDTFESMPNBEGRYTFGASCVA 298
Db 240 HCAGCTGPRATDCLACRDFNDGTCKDTPCKPIKIDYISHGVDPNPKTYTFGAACVRE 299
Qy 229 CRYNYLSTDVSGCTVLCPLHNOEVTAEPTORCEKSKCARCVGLGMEHLREVAVTS 358
Db 300 CFSNTVYVE--GACVNSCSAGMLEVD--ENGRSKCPQDGVCPKVCDDIGSGLSNTIAVNS 357
Qy 359 ANIOEPAGCKKIFGSLAFLEPESFGDPASNTAPLOEOLQVETLEITGYLYISAMPSS 418
Db 356 TMRBPSNTKINGDILNRSNESPDPHKTIGTMDPEHMLNLTVEITGYLYIVMMPEN 417
Qy 419 LPLDSVFNQVIRGRILHNGAYS--LTLQGLISWLGRLSLRELSGGLALIHNTLCEV 477
Db 418 MTSLSVFNQVIRGRITTSRSGFSFVYVQVRHLQWLGRLSLRELSGGLALIHNTLCEV 477
Qy 478 HTPVPDQVLRNPHQALHTANRBEDECEGLACHQLCARGHCWGBPTQCVNCSGFLAG 537
Db 478 NTINNRRLRFRSDQSEYDART-----ENQTCNECSBDCGW--PGPTMCVSCILHVDG 529
Qy 538 QCEVSECRVLOGLPREYVNAHCLPCHPECPONGSVTCFGEPAQDCVCAHYKDDPPCY 597
Db 530 GRCVASCNLQGBRRAQVDGRVCQCHQCVLTQDSTLTCTGPRPANCSSAHRQDDPQCI 589
Qy 558 ARCPGVKPDLSYMPIWKPDPDEBACQPCPINCTHSQVLDLDDKCPAEGPASPISIIA 657
Db 590 PRCPFGIILDDGDTL--IMKYADKMGQCPQCHQNTQCSGPGLSGCGD--IVSHSLAVGL 647
Qy 658 VGLILLVAVLVGVLIRBROOKRKMTMRBLQTELYEPLTPSGAMPNQOMKILE 717
Db 648 VSGLLITVAVLLVLLRRRRK--RKRTIRCLQKELEVEPLTPSGQAPNQAFILKE 706
Qy 718 TELRKVYVGSAGFTVYKGIWIPDGENYKIPVALIYLRNTPSKANKSILDEAYMAGV 777
Db 707 TEPKQDRLVGSAGFTVYKGIWIPDGENYKIPVALIYLRNTPSKANKSILDEAYMAGV 766
Qy 778 GSPYVSRLLIGLCTSTQVLTQMLPGCLLDHVRENRGLSGQDLNMCQIANKGNSYLE 837
Db 767 DHPHCRLLIGLCTSTQVLTQMLPGCLLDHVRENRGLSGQDLNMCQIANKGNSYLE 826
Qy 838 DVLVYHRDLAARVLYKSPHYKTTDPGLARLLDIDETRYHADGKVPYIKMALBSILR 897
Db 827 ERLVYHRDLAARVLYKSPHYKTTDPGLARLLDIDETRYHADGKVPYIKMALBSILR 886
Qy 898 RPTHOSDVWSYGVYTWELMTFGAKPYDGLPAREIPDLKRGRLPOPICTIDVYVIMYK 957
Db 887 TYTHOSDVWSYGVYTWELMTFGSKPYDGLPAKEIASLVEGERLPPPICTIEVYVIMYK 946
Qy 958 CMMDISECRPRRELIVSESRMARDPQRFVIONEDLGPASPLDSTFYRSLLEDDMDGL 1017
Db 947 CMMDISECRPRRELIVSESRMARDPQRFVIONEDLGPASPLDSTFYRSLLEDDMDGL 1001
Qy 1018 VDAEYLVQGFPCPDPAAGAGVHHRRSSSTSGGDLTLGLEPSEBBAAPSLAP 1077
Db 1002 VDAEYLVQGFPCPDPAAGAGVHHRRSSSTSGGDLTLGLEPSEBBAAPSLAP 1024
Qy 1078 SRGASDVVDGDLGMAKGLQSLPTHDPSPLORYBEDPTV--PLSETDGYVAPLCSQ 1136
Db 1025 PFGH-----PVRENSITLRNIDPTONALEKDLQGH----- 1055
Qy 1137 PEYVQPPVAPOP-----PSPRE-----GFLP--AARPAATLBRPKTSLSPKNGVVD 1183
Db 1056 --EYVQPPSETSRSLSDINPNVEDLTDGMPVSSQGBAETNFSRPEYVNTQNL----- 1111
Qy 1184 VPAFGAVENPEYLTQGGAAPOHPPPAPAFDNLVYWDODPBERGAPSTFKCTPTA 1243
Db 1112 PLVSSGSKMDPPY--QAG-----YQAAF-----LPQTGALTGNGMFLPAA 1149

```

```

Qy 1244 ENPEYLG 1250
Db 1150 ENLEYLEG 1156

RESULT 9
A36223
kinase-related transforming protein (erbB3) (EC 2.7.1.-) precursor - human
C.Species: Homo sapiens (man)
C.Date: 04-Oct-1991 #sequence_revision 13-Jan-1993 #text_change 17-Nov-2000
C.Accession: A36223; 159164
R.Kraus, M.H.; Issing, W.; Miki, T.; Popescu, N.C.; Aaronson, S.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 9193-9197, 1989
A.Title: Isolation and characterization of ERBB3, a third member of the ERBB/epidermal
A.Reference number: A36223; MUID:90083234; PMID:2687875
A.Accession: A36223
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-1342 <KRA>
A.Cross-references: GB:M29366
R.Plowman, G.D.; Whitney, G.S.; Neubauer, M.G.; Green, J.M.; McDonald, V.L.; Todaro, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990
A.Title: Molecular cloning and expression of another epidermal growth factor receptor-
A.Reference number: 159164; MUID:90311312; PMID:2164210
A.Accession: 159164
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-559, 'G', 561-957, 'F', 959-1063, 'G', 1065-1342 <RSS>
A.Cross-references: GB:M34303; MUID:9183990; PIDN:AAA35979.1; PID:9306841
C.Gene: ERBB3; HER3
A.Gene: GDB:ERBB3; HER3
A.Cross-references: GDB:119880; OMIM:190151
A.Map position: 12q13-12q13
C.Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
C.Keywords: ATP; phosphotransferase
F:707-972/Domain: protein kinase homology <KIN>
F:715-723/Region: protein kinase ATP-binding motif

Query Match 35.7%; Score 2432.5; DB 2; Length 1342;
Best Local Similarity 40.7%; Pred. No. 4,5e-92;
Matches 533; Conservative 191; Mismatches 458; Indels 129; Gaps 32;

Qy 10 GILLALPPGAA--STOVCTGTDMKRLPASBETHLDMRLHYOGQGVVQGNLELYLPT 67
Db 11 GILFSLARSEVNGQAVCPGTLNGLSVGDENQVOTLYKLYEREVEVWNGNLEIYLTGH 70
Qy 68 NALSLEFLQDIOGVGYLLAHNQVQVPLQRLIRYGTQLFEDVYALAVLDNGDPLNNTT 127
Db 71 NADLSFLQWIRREVYGVYLVAMNBFSTLPLNLRVVGTVGQYVDGKFAIFVM-----LNYNT 125
Qy 128 PYTGASFGRLRLQRLSLTEILLKGVLIQNRNQLCYQDITLIMWDIHHKNNQLALLTLDTN 187
Db 126 ---NSSHALRQLRLQLTEILLSGVYIEKNDLCMDITLDMWDIVRDR--AEIVVKD 178
Qy 188 RPARCPGSPCKSGSRGWESSBDCQSLTRTVCAAGC--ARCKPLPTDCHEOCAAGCTG 246
Db 179 NGRSCPCHVEVKG--KCMPSGSDCQTLTKITCAPCQNGHCFBPNNQCHDCAGGCSG 237
Qy 247 PHSDDLACLFNHSIGELHCPALVTYNTDTFESMPNBEGRYTFGASCYVACPYNVLT 306
Db 238 PDLTDFACRHNNDNSACVPRCPQVLYKGLFQLBEPNHTKQYGVGVCAACPFHNV--V 296
Qy 307 DYGSCTVCPRLANQVTAEDGTQRCCKSKPCARVYCYGLGMHLREVRATVANSNIOFAG 366
Db 297 DQTSVCRACPDPKMEVD--KNGKMCBPCCGLCPKACBEGSG--SRFQTDSSNIDGFVN 353
Qy 367 CKKIFGSLAFLEPESFGDPASNTAPLOEOLQVETLEITGYLYISAMPDSLPDLSVFQ 426
Db 354 CTKLIGNDLFTTGLNGDPWHKIPALDPEKLVNFRVREITGLNIOSNRPHHNPVSFS 413
Qy 427 NQVIRGRILHNGAYS--LTLQGLISWLGRLSLRELSGGLALIHNTLCEVATVPWDOL 485

```

```

Db      414 NLTTIGRSILYNRGFSLLIMKNTVTSIGFRSLKEISAGRIYISAKRQLCTHHSIMTKV 473
Qy      486 FRNPQALLHTA-NRPEDECVGEGLAGCHOLCARGCHWGPPTQCVNCSQPLRQOECEVEC 544
Db      474 LRQPTREERLDIKNNRPRARDCAVAGKVCIDPLCSSGCGPFGQCLSCRYN SRGGVCTYHC 533
Qy      545 RVLGGLPREVYVNAHCLPCHPECOPONGSVTCFGRPADCVACAHYDPPFCVACRCSGV 604
Db      534 NPLNGSPREBAHAECFSCHECOPMGTATCNGSGSDTCAQCHFRDGHCVSSCHGV 593
Qy      605 KPLDSWMPKIPDEBEGACOPCPINCHSGVDLDDKCCPAEORA---SPLTISIAVVG 660
Db      594 LG--AGPIKYKVDVQNECRPHENCTQCKGPELDQCLQTLVLKTKTLTALTYIAG 651
Qy      661 ILLVVVLGVVFGILIKRROKIR-KYTMRRLLQETELVEBPLTSGAMPNOAKRILKETE 719
Db      652 --LVVIFMMLGTFVLVYRGRIQNKRAMRYLERGESIEPLDPS-BKANKVLAIRFETE 708
Qy      720 LRKVKTLGSGARCTVYKGIWIPDGENVKIPVALKYLBENTSPKANKELDEAYVMAGVS 779
Db      709 LRRLKVLGSGVFTVHKGVMIPEGESIKIPVCIKVIDKSGROSFOAVTDHMLAIGSLDH 768
Qy      780 PYVSRLLIGICTVTLVQVLTQMPYGLLDHYRBNRGLSGODLLNMCQIAKGMSTYLEDV 839
Db      769 AMLVRLIGLCPGSSLDLVLYPLGSLLDHYRQRRGALGPOLLNMGVQIAKGMSTYLEDV 828
Qy      840 RLVRRLDAANVLVKSFNHYKITDPLARLLDIDETRYHADGGKVPKIMNALSILRRF 899
Db      829 GNVHRLAANVNLKSPSQVQVADPFGVADLLPPDDQLVLSBAKPTIKMALLSIHKGX 888
Qy      900 THOSDWSYVYTWELMTFGAKPYDGIIPABEIPDLKESGRLPORPCTIDVYIMVKCM 959
Db      889 THOSDWSYVYTWELMTFGAEPPYAGRLAEVPLLEKGRLLAOPCTIDVYIMVKCM 948
Qy      960 MIDSECRPRELVSEPRARDPORFVNIQNEIDLGA---SPLDSFYFSLLEDDMGD 1016
Db      949 MIDBNRPTKELANETTRARDPRLVLIKRES-GEGLAPGEPHGLTKKILBEVLEP 1007
Qy      1017 LVDAEYLVPQGFPCDPAPAGAGVNHRRSSSTRSGGDLTLGLEP-SEEBAPRSP 1075
Db      1008 ELRLDLDAEED-----NATTTLSALSLPVGTILNRPGRGSL 1048
Qy      1076 APBEGAGSDVFDGDLGMAAKGLOSLPTHN-PSPLQKYSDDPYVLEP-----SEIDGYV 1128
Db      1049 SPBSGY-MPNQGNLGBSCQESAVSGSSEKCPREVSLH-----PMERGCLASBESSEGHV 1101
Qy      1129 A-----PLTCSPOPE---YVNPQDVPRPSPRSGP-----L 1157
Db      1102 TGSEBALOEKVSICRSRKRSPRPDDSAVHSQHSLLTPVPLSPGLSEEDVNGYV 1161
Qy      1158 PAARPAATLERPKTLEP-GRNGV-----KQVAFAGAVENPEYLTPOGGAAPQPHP 1210
Db      1162 PDLHNGTSPBSRGSTLSSVGLSGLTEBEDED-----EEVEYNNRRRBSHP-PHP 1212
Qy      1211 PARSAPADNLYND-----QDPRKCAPSTFKGTPLAKPEYTL 1249
Db      1213 RPSLSELGELYNDVGSGLSASLGSTQSCPLHPVPIPTAGTTPDEDEYEVY 1263

```

RESULT 10

```

epidermal growth factor receptor homolog precursor - rat
N/Alternate names: ErbB3 protein; HER3 protein
C/Species: Rattus norvegicus (Norway rat)
C/Date: 17-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 13-Nov-1998
C/Accession: J04387
Gene 165, 279-284, 1995
A/Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protein.
A/Reference number: J04387; MUID:96096535; PMID:8522190
A/Accession: J04387
A/Molecule type: mRNA
A/Residues: 1-1339 <HEL>

```

```

A;Cross-references: GB:U09339; NID:9915389; PID:9915390
A;Experimental source: Liver
A;Note: The authors translated the codon AAC for residue 369 as Thr and GTT for residue
C;Comment: This protein is a functional heregulin receptor that transduces signals to ti
C;Gene: ErbB3
C;Superfamily: Unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-139/Product: epidermal growth factor homolog #status predicted <MAT>
F;640-659/Domain: transmembrane #status predicted <TM>
F;705-970/Domain: protein kinase homology <KIN>
F;713-721/Region: protein kinase ATP-binding motif
F;939,1051,1156,1196,1219,1257,1259,1286,1325/Binding site: phosphate (Tyr) (
Query Match 34.4%; Score 2347.5; DB 2; Length 1339;
Best local Similarity 40.8%; Pred. No. 1,3e-88;
Matches 524; Conservative 170; Mismatches 434; Indels 155; Gaps 34;
Qy      3 LNALCRWGLLALLPRGA--STQVCTGDMKRLRPASEETHLDMRLHYQCGQVQGN 59
Db      7 LQVLC---FLSLARGSEKNSQAVCPGLNGLSTGADNQGQVLYKLYKECEVMGN 62
Qy      60 LELTYLPTASLSFLDIOEGVQVYLIANNQVQVPLQRLRYRGQLFEDNVALAYLDN 119
Db      63 LELVTLGNHADSFLQWIREVTAYVLANNEFSVLPPLNRYVVGQVQDGKFAIYV-- 120
Qy      120 GDLPLNNTPTVATGSPGLRELOLRSLTEILKSGVLIQBNPLCYQDPTILMKDIFHNKQL 179
Db      121 --LNTNT-----NSIALNQLKFTQLTLEISGVYIEKNDKLCMDITIDRDIYVR-- 170
Qy      180 ALTLIDTNSRACHPCSPMCKSGRCWGSSESDCQSLTRIVCAGG-ARCKGRLPTDCHE 238
Db      171 GAEIVKNNGANPCPCHGVCKG-RCWGPGRDDCQILTKTICAPQCNCRCPGPNQCCHD 229
Qy      239 QGAAGCTGPRHSCTACAFHNHSGICEHLNCPALVTYNTDFEEMPNREGVYTRASCTA 298
Db      230 EGAGSGSGQDIDCFACRRFNSGACVPRCPRELVNKLTFQLEPNHTTYOYGCVVAS 289
Qy      299 CFYNYLSTVDSCTVLCPLHNOEVTADGTQRCSEKSKPCARVCYGLMEHLREVRVTS 358
Db      290 CPHNFP-VQDPTVRAACPRDKKEVD-KHGLAKMEPCGGLCPKACEGSG--SRVQVDS 345
Qy      359 ANIQEAGACKIFGSLAFPESEFGDPASNTADLOEQLQVFTLEETGYLYISAMPDS 418
Db      346 SNIDGFVNTCKIIGNDLITGLAVDPBMHKIPALDEKIANVFVREITGYLTIQSWPH 405
Qy      419 LPLSLFQNLQVIRKILNKGAYS-LTQGLGISWGLRSLRELGSLALIHNTILCFV 477
Db      406 MNPVSFNSLTTTGGSLYNRGFSLLIMKNTVTSIGFRSLKEISAGRYISANQQLCYH 465
Qy      478 HTVPQDLFRNPQALLHTA-NRPEDECVGEGLAGCHOLCARGCHWGPPTQCVNCSQPLR 536
Db      466 HSLMTRLRKGSEERLDIKYRPLRGLGCLABGVCPDLCSGGCGWGAPOCLSCRYNKR 525
Qy      537 GQECVEECRYLGLPREYVNAHCLPCHPECOPONGSVTCFGRPADCVACAHYKDRPFC 596
Db      526 EGVCTVHNCNPLQGEPRFVHAQCFSCHEPCLPMBGISTYNGSGSDACACAHFRDPSHC 585
Qy      597 VARCPSGVKPDLSYMTMKPDEBEGACOPCPINCHTSC--VDDDKCPAEQARSPLTSI 654
Db      586 VNSCPHGIIG--AKGPIKYKPDQONCRPHENCTQCKGPELDQCLQGEAEVLSKPHLV 643
Qy      655 ISAVGILLVVVLGVVFGILIKRROKIR-KYTMRRLLQETELVEBPLTSGAMPNOAKMR 713
Db      644 IAYTVG--LAVILMIGSFLYKRGRIQNKRAMRYLERGESIEPLDPS-BKANKVLAIR 700
Qy      714 ILKTELKRVKVLGSGAFGVYKGIWIPDGENVKIPVALKYLBENTSPKANKELDEAYV 773
Db      701 IFKTELKRLKVLGSGVFTVHKGVMIPEGESIKIPVCIKVIDKSGROSFOAVTDHMLA 760
Qy      774 MGVGSPYVSRLLIGICTVTLVQVLTQMPYGLLDHYRBNRGLSGODLLNMCQIAKGM 833

```

Db 761 VGSLLAHAVRLGLCGSSLDLVTQYLPGLSLDHVKHRETLGQOLLINMGVOIAKGM 820
 Qy 834 SYLEDVRLVHRDLAAANVLKSPNNHYKITDFGLARLLIDETEVHADGKVPITKMALES 893
 Db 821 YLIEBHSVMHRDLALNNVLMKSPSQVADLPPDDKQLHSEAKTPIKMMALLES 880
 Qy 894 ILRRRTTHOSDWSYGVTTWELMTFGAKPYDGI PAEIPDLLEKGERLPQPICTIDVYM 953
 Db 881 IHFGKTHOSDWSYGVTTWELMTFGAPYAGLRLEIPDLLEKGERLAPQICTIDVYM 940
 Qy 954 IAWKCMIDSECPREELVSEFSRMAADPQRFVUQNEDELGPASPLDSFFYSLLIEDD 1013
 Db 941 VAWKCMIDENIRPTFEELANETRMARDPRLVIGAS-GEQTP--PAAPSVLTKE 997
 Qy 1014 MGDVLAEYLVPOGFCPPAPGAGVNHHRSSSTRSGGGLTLGLEPSEE----- 1068
 Db 998 L-----QEALEBEL-----DLDLDLEAEEGLATS 1023
 Qy 1069 -----EAPNSPLAPSEB-----AGSDVFDGLGMAKGLQSLPTHD 1105
 Db 1024 LGSALSLPTGTLTRPRGSGSLSPSSGYMPMNSLGEACLDASAVLGREGQFSPISLH- 1082
 Qy 1106 PPSLQVSEDPYPLPSFTDGYV---APL-----TC-----SPQPE---YVNOQPDV 1145
 Db 1083 PLPRGR-----PASESSEGHVTGSEAELOEKVSVCRSRSPRPRGDSAYHSQRHS 1135
 Qy 1146 RPQPSPRSGP-----LPAAAPAGATLERPKTILSP-GKNQVY-----KQVFAP 1187
 Db 1136 LPLTPVPLSPGLSEBDGNGVWPDTHLRGASSRSGTLLSVGLSTGLTEEDBD----- 1191
 Qy 1188 GGAIVENPEYLTPOGGAAPORPP 1210
 Db 1192 ---BEVEYNNRRKRGSP-PRPP 1209

RESULT 11

TVFVLV

protein-tyrosine kinase (EC 2.7.1.112) erbB - avian leukosis virus

N:Contains: amino end of gag protein; env protein fragment; protein-tyrosine kinase

C:Species: avian leukosis virus, ALV

C>Date: 31-Dec-1991 #sequence_rev150 31-Dec-1991 #text_change 11-Jun-1999

C/Accession: B00643; A00643

A/Accession: B00643

A/Molecule type: mRNA

A/Residues: 1-698 <N1>

A/Cross-references: GB:M10066; GB:M13881; NID:G211749; PIDN:AAA48763.1; PID:G211750

A/Note: in Genbank entry CHKBRBF, release 109.0, the source is designated as Gallus gall

C/Comment: This protein is synthesized as a gag-env-erbB protein.

C/Genetics:

A:Gene: gag-env-erbB

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C/Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific P

F.1.6/Product: gag protein (fragment) #status predicted <GAG>

F.7.59/Product: env protein (fragment) #status predicted <ENV>

F.60-698/Product: protein-tyrosine kinase erbB #status predicted <ERB>

F.194-459/Domain: protein kinase homology <KIN>

F.202-210/Region: protein kinase ATP-binding motif

F.229/Active site: Lys #status predicted

Query Match 25.94; Score 1765.5; DB 1; Length 698;
 Best Local Similarity 52.24; Pred. No. 3.5e-65;
 Matches 374; Conservative 80; Mismatches 137; Indels 125; Gaps 18;

Qy 578 GPRADOCVACAKYDPFPCVACPSGVKPDLSYVIMKPPDEBACOPCPINCHSVDL 637
 Db 60 GP--DRCMKAMRIDGHCYKACPAAGVLSGVDLT-VKIVADANAACVCLPNCIRGSKSP 116
 Qy 638 DDKGCPAQRASPLTISAVV-GILLVVVGAVFGILRRQOKIRKTYMRLLQETEL 696

Db 117 GLEGGP---NGSKTSPIAAGVGGILCLVYVVGIGIGLYLRR-HIVKRTLRRLDEREL 172
 Qy 697 VEPPLTSGAMPNOQMRILKETELARKVYLGSGAFGVYKGIWIPQENYKIPVALKVL 756
 Db 173 VEPPLTSGAMPNOAHRIILKETELPKKVKYLGSGAFGVYKGIWIPSEKXKIPVALIELR 232
 Qy 757 EHTSPANKELIDEAVYVAGVSPVYSLIGTILSTVOLVTLQMPYGCILDHVRENRGR 816
 Db 233 EHTSPANKELIDEAVYVAGVSPVYSLIGTILSTVOLVTLQMPYGCILDHVRENRGR 292
 Qy 817 LSSODLLNMOIAKMSVLEDEVLVHRDLAAANVLKSPNNHYKITDFGLARLLIDETEV 876
 Db 293 IGSQYLLNMCVOIAKMSVLEDEVLVHRDLAAANVLKSPNNHYKITDFGLARLLIDETEV 352
 Qy 877 YADGKVPITKMALESILRRRTTHOSDWSYGVTTWELMTFGAKPYDGI PAEIPDLLE 936
 Db 353 YABEGKVPITKMALESILHRIYTHOSDWSYGVTTWELMTFGSKPYDGI PAEISSVLE 412
 Qy 937 KEERLPQPICTIDVYIMVWCMIDSECPREELVSEFSRMAADPQRFVUQNEDELG 995
 Db 413 KEERLPQPICTIDVYIMVWCMIDADSRPKRELIAPERSKMARDPRLVIGQDERRM 472
 Qy 996 PPSLQVSEDPYPLPSFTDGYV---APL-----TC-----SPQPE---YVNOQPDV 1055
 Db 473 LPSPTSKPRTLTMESEDMEDIVDAEYLVPOGFCPPAPGAGVNHHRSSSTRSG 513
 Qy 1056 GEDLTGLLEPSEBEAPRSP-----APSEAGSDVFDGLGMAKGLQSLPTHDPSPLQ 1110
 Db 514 -----SRTLPLSLSATSNNSANNCID-----RNGQGHPRAREDSFVQ 550
 Qy 1111 RYSEDPVPLPSET--DGVAAPITGSPQRYVNOQPDVPRPQPSRSGPLPAAPAGATLE 1168
 Db 551 RYSEDPVPLPSET--DGVAAPITGSPQRYVNOQPDVPRPQPSRSGPLPAAPAGATLE 585
 Qy 1169 RPKTILSPGKNQVYKQVF-----AFGAVENPEYLTPOGGAAPORPPPAF 1213
 Db 586 ---TAMVONQIYNNISLTAISKLPMDSKYONSHSTAVDPEYL-----NTNQSPLA 633
 Qy 1214 SFPEFNLVYWDQ-----DPEE-----RGAPSTFKGTPTAENPEYLGIDVP 1254
 Db 634 KTVFESSPFWIGSGNHQINLNDPQODPLPNTKXNGLKVAHAENPEYLRVAAP 689

RESULT 12

TVYUHL

protein-tyrosine kinase (EC 2.7.1.112) erbB - avian erythroblastosis virus (strain H)

C:Species: avian erythroblastosis virus

C/Date: 18-Apr-1984 #sequence_rev150 18-Apr-1984 #text_change 11-Jun-1999

C/Accession: A00644; A38022

R/Yamamoto, T.; Nishida, T.; Miyajima, N.; Kawai, S.; Ooi, T.; Toyoshima, K.

Cell 35, 71-78, 1983

A/Title: The erbB gene of avian erythroblastosis virus is a member of the src gene fami

A/Reference number: A00644; MUID:84026539; PMID:6133229

A/Molecule type: DNA

A/Residues: 1-604 <YAM>

A/Cross-references: GB:K01216; NID:G209676; PIDN:AAA42400.1; PID:G209678

R/Debnire, B.; Henry, C.; Benaisa, M.; Biserte, G.; Claverie, J.M.; Saule, S.; Martin,

Science 224, 1436-1439, 1984

A/Title: Sequencing the erbA gene of avian erythroblastosis virus reveals a new type of

A/Reference number: A38022; MUID:84223957; PMID:6328658

A/Molecule type: DNA

A/Residues: 1-26 'W', 30-119 'F', 141-145 'V', 147-152 <DEB>

A/Cross-references: GB:K02006

C/Genetics:

A:Gene: erbB

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C/Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific I

F.130-395/Domain: protein kinase homology <KIN>

F.138-146/Region: protein kinase ATP-binding motif

F.165/Active site: Lys #status predicted


```

Db      715 QYTAGPY-----CAASPRSSKLTAND-----VNMIFITGAVLEPTIC 755
Qy      669 VVFGI-LIKRROOKIRKKT--MRLLLOETELVEPLTPSGAMPQAOQRILKETELRKYV 725
Db      756 ILCVYVYICRQKAKKETVKTMTALSGRDSRPLRPSNIGALCKLRKYDAELKRGV 815
Qy      726 LSGAGFTVYKGIWIDGENVKI PVAIKVLRNTSPANKREILDEAVYMAVGSPPVYSRL 785
Db      816 LGMGAGRVYKGVWVEGENVKI PVAIKELKSTGASSSEPLREAVYMASBEHVNLKL 875
Qy      786 LGLCLSTVQVYQVLMPLGCLLDHYRNRGRSLGSDLLANCMQIAGMSTLEDRVLVHRD 845
Db      876 LAVGSSQWMLITQMLPLGCLLDYRNRNRDKISKLLMNSTQIAKMSYLEEKRLVHRD 935
Qy      846 LAARNVLYK---SPNVKICTDGLARLLIDETERYHADGKVPVKMMALSLIRRPETHQ 902
Db      936 LAARNVLYKLAGEDH---DFGLAKLLSDSNRYRAAGKMPVKMLALECIRNRYPTSK 991
Qy      903 SDVMSYGVTVWELMTFGAKPYDGI PPAEIPDLLEKGERLPQPICTIDVYIMVYKCMMD 962
Db      992 SDVMSYGVTVWELMTFGAKPYDGI PPAEIPDLLEKGERLPQPICTIDVYIMVYKCMMD 1051
Qy      963 SECRPRELVSEFSRMAADPQRFVYIQNEDLG--PASPLDSTPTYSLEDD--DMGDL 1017
Db      1052 AAMRPFPKQLTIVFAFARPPRYLAII LGGKFTRLPA-----YTSQDEKDIIRKLAPT 1104
Qy      1018 VAABEVLVPOQGFPCDDPAPGAGCVHHRSSSTRSGGDLTLGLEPSEEARP----- 1071
Db      1105 TGQSEAIAPRDYLOKRALGPS-----HRTDCT-----DENPKLNRVC 1143
Qy      1072 RSPPLASBEGAGSDVFDG--DLGMAKAGLQSLPTHDPBLORYSBEPVPLPSETDGYV 1128
Db      1144 KQPSNKGSSSTGDERDSAREVGVNLR-----LDLPVEDDDYL 1182
Qy      1129 APLTCSPOPEYVQPDVRPPSPRSGPLPAAPAGATLERPRLTSKGVKGVKDVAFAG 1188
Db      1183 MP-TCPGPPNNNNNM-----NPNQNMMAVGAAGYM-----DLIGVP 1220
Qy      1189 GAVENDEYL-----TPGGAAPOPH-----PPAFSP-APDNLYYWD 1224
Db      1221 VSVQNEPRLNAGTLGVGESPIPTQITIGVWGPGPMGVKVPMPGSEPTSDHEIYND 1279

```

RESULT 14

S35745
protein-tyrosine kinase (EC 2.7.1.112) erbB - avian erythroblastosis virus
C/Species: avian erythroblastosis virus
C/Date: 03-Mar-1994 #sequence_revision 26-May-1995 #text_change 28-Feb-1997
C/Accession: S35745
R/Vennstrom, B.
submitted to the EMBL Data Library, March 1993
A/Reference number: S35743
A/Molecule type: DNA
A/Residues: 1-544 <VEN>
A/Cross-references: EMBL:X12707
C/Genetics:
A/Gene: erbB
C/Superfamily: epidermal growth factor receptor; protein kinase homology
C/Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific P
F;135-400/Domain: protein kinase homology <KIN>
F;143-151/Region: protein kinase ATP-binding motif
F;170/Active site: Lys #status predicted

Query Match 24.1%; Score 1645; DB 2; Length 544;
Best Local Similarity 57.7%; Pred. No. 2.3e-60;

Matches 335; Conservative 65; Mismatches 97; Indels 84; Gaps 13;

```

Qy      578 GPADQVCAAHKDPFCVACPSGKPLSTYMPKRPDEBACQPCPINTCHSCVDL 637
Db      1 GP--DHCMKCAHIDGPHCVKACPAVGLGENDTL-VKRYADANAVCQLCHPNCIRGCKGP 57

```

```

Qy      638 DKGCPAEGAPASPLTIIISAV-GIILVVYLVGVFGILIRROOKIRKTMRLLOETEL 696
Db      58 GLEGCP---NGSKTPIAAGVGLCLLVVVGIGIYLERR-HIVKRTLRLLOEREL 113
Qy      697 VEPLTPSGAMPQAOQRILKETELRKYVYKSGAFGVYVYIMPODENVKI PVAIKVLR 756
Db      114 VEPLTPSGAMPQAOQRILKETELRKYVYKSGAFGVYVYIMPODENVKI PVAIKVLR 173
Qy      757 ENTPSPANKREILDEAVYMAVGSPPVYSRLGICLTSTVQVLTQVLMPLGCLLDHYRNRGR 816
Db      174 EATSPANKREILDEAVYMASVQNPVHVCRLGLICLTSTVQVLTQVLMPLGCLLDHYRNRGR 233
Qy      817 LGSQDILLNCMQIAKMSYLEDVLYHRDIAANVLYKSNHYKTYDGLARLLIDETE 876
Db      234 IGSQYLLNMCVQIAKMSYLEDVLYHRDIAANVLYKSNHYKTYDGLARLLIDETE 293
Qy      877 YEADGKVPVKMMALSLIRRPETHQSDVMSYGVTVWELMTFGAKPYDGI PPAEIPDLLE 936
Db      294 YEADGKVPVKMMALSLIRRPETHQSDVMSYGVTVWELMTFGAKPYDGI PPAEIPDLLE 353
Qy      937 KGERLPQPICTIDVYIMVYKCMMDSECRPRELVSEFSRMAADPQRFVYIQ-NEPDG 995
Db      354 KGERLPQPICTIDVYIMVYKCMMDSDASRPRERELAESKMAADPRLVYIGDERMH 413
Qy      996 PASPLDSTPTYSLEDDMGDLVDAEYLVPOQGFPCDDPAPGAGCVHHRSSSTRSG 1055
Db      414 LPSPTSKFYRTLMEEDEMDIADAEYLVPHQGF-----NSPST--- 454
Qy      1056 GBDULTIGLEPSEEARSP-----APSEGAGSDVFDGDLGMAKAGLQSLPTHDPSPLO 1110
Db      455 -----SRPLSLSLATSNNSATNCIDRNG-----H----- 481
Qy      1111 RYSEDPVPLPSETDGYVAPLTCSPOPEYVQPDVRPPSPS 1151
Db      482 -----PVREDGFL-----PAPEYVQ--LMPKKS 504

```

RESULT 15

S00727
kinase-related transforming protein (erbB) (EC 2.7.1.1.-) - avian erythroblastosis virus
C/Species: avian erythroblastosis virus
C/Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 28-Feb-1997
C/Accession: S00727
R/Scotling, P.; Vennstrom, B.; Jansen, M.; Graf, T.; Beug, H.; Hayman, M.J.
Oncogene Res. 1, 265-278, 1987
A/Title: Common site of mutation in the erbB gene of avian erythroblastosis virus mutan
A/Reference number: S00727; PMID:88217326; PMID:2897102
A/Accession: S00727
A/Molecule type: DNA
A/Residues: 1-545 <SCO>
A/Cross-references: EMBL:X06943
C/Genetics:
A/Gene: erbB
C/Superfamily: epidermal growth factor receptor; protein kinase homology
C/Keywords: ATP; phosphotransferase
F;135-400/Domain: protein kinase homology <KIN>
F;143-151/Region: protein kinase ATP-binding motif

Query Match 24.0%; Score 1638; DB 2; Length 545;
Best Local Similarity 57.7%; Pred. No. 4.4e-60;

Matches 335; Conservative 64; Mismatches 98; Indels 84; Gaps 13;

```

Qy      578 GPADQVCAAHKDPFCVACPSGKPLSTYMPKRPDEBACQPCPINTCHSCVDL 637
Db      1 GP--DHCMKCAHIDGPHCVKACPAVGLGENDTL-VKRYADANAVCQLCHPNCIRGCKGP 57
Qy      638 DDGCPAEGAPASPLTIIISAV-GIILVVYLVGVFGILIRROOKIRKTMRLLOETEL 696
Db      58 GLEGCP---NGSKTPIAAGVGLCLLVVVGIGIYLERR-HIVKRTLRLLOEREL 113
Qy      697 VEPLTPSGAMPQAOQRILKETELRKYVYKSGAFGVYVYIMPODENVKI PVAIKVLR 756
Db      114 VEPLTPSGAMPQAOQRILKETELRKYVYKSGAFGVYVYIMPODENVKI PVAIKVLR 173

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 16:06:24 ; Search time 18 Seconds
(without alignments)
3278.806 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815
Sequence: 1 MELALCRWGLLALLPPGA.....TFKGTPTANPEYLGLDVPV 1255

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by change to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	1 ERB2_HUMAN	P04626 homo sapien
2	6003	88.1	1257	1 ERB2_RAT	P06494 rattus norv
3	5993.5	87.9	1254	1 ERB2_MESAU	Q60833 mesocricetu
4	3167	46.5	1210	1 EGFR_HUMAN	P00533 homo sapien
5	3146	46.2	1210	1 EGFR_MOUSE	Q01279 mus musculu
6	3004.5	44.1	1308	1 ERB4_HUMAN	Q15303 homo sapien
7	2989	43.9	1308	1 ERB4_RAT	Q62956 rattus norv
8	2724.5	40.0	1167	1 XMRK_XIPMA	P13388 xiphophorus
9	2441.5	35.8	1332	1 ERB3_HUMAN	P21860 homo sapien
10	2370.5	34.8	1339	1 ERB3_RAT	Q62799 rattus norv
11	1975	29.0	1426	1 EGFR_DROME	P04412 drosophila
12	1748.5	25.7	634	1 ERBB_ALV	P00534 avian leuko
13	1702	25.0	604	1 ERBB_AVIER	P00535 avian eryth
14	1628	23.9	540	1 ERBB_AITIU	P11373 avian eryth
15	1623	23.8	703	1 EGFR_CHICK	P13387 gallus gall
16	1301	19.1	1367	1 LTR3_CAERL	P24348 caenorhabdi
17	1142.5	16.8	245	1 ERB2_MOUSE	P70424 mus musculu
18	736	10.8	1363	1 ILPR_BRALA	O02466 brachiocto
19	715	10.5	1382	1 INSR_HUMAN	P06213 homo sapien
20	710	10.4	1383	1 INSR_RAT	P15127 rattus norv
21	709.5	10.4	1372	1 INSR_MOUSE	P15208 mus musculu
22	706	10.4	1300	1 IRR_MOUSE	Q9W614 mus musculu
23	703.5	10.3	1607	1 MIRP_LYMER	Q25410 lymphoma ara
24	699	10.3	1297	1 IRR_HUMAN	P14616 homo sapien
25	694.5	10.2	1300	1 IRR_CAVPO	P14617 cavia porce
26	682.5	10.0	1477	1 HTR7_HYDAT	Q25197 hydat atten
27	651	9.6	1367	1 EGFR_HUMAN	P00869 homo sapien
28	643	9.4	1373	1 IGIR_MOUSE	Q60751 mus musculu
29	639.5	9.4	1370	1 IGIR_RAT	P24062 rattus norv
30	627	9.2	1390	1 INSR_AEDAE	Q93105 aedes aegypt
31	622	9.1	2146	1 INSR_DROME	P09208 drosophila
32	607	8.9	987	1 EPB4_HUMAN	P54760 homo sapien
33	590.5	8.7	984	1 EPB1_CHICK	O07944 gallus gall

34	589.5	8.7	977	1 EPB2_MOUSE	Q03145 mus musculu
35	587	8.6	1114	1 RET_HUMAN	P07949 homo sapien
36	585	8.6	987	1 EPB4_MOUSE	P54761 mus musculu
37	584.5	8.6	976	1 EPB2_HUMAN	P29317 homo sapien
38	582.5	8.5	984	1 EPB1_RAT	P09759 rattus norv
39	578	8.5	902	1 EPBB_XENLA	Q91736 xenopus lae
40	576.5	8.5	984	1 EPB1_HUMAN	P54762 homo sapien
41	576.5	8.5	985	1 EPBA_XENLA	Q91571 xenopus lae
42	575	8.4	1053	1 PAK1_CHICK	Q00944 gallus gall
43	569	8.3	1058	1 PAK1_XENLA	Q91738 xenopus lae
44	567	8.3	757	1 HTR16_HYDAT	P53356 hydat atten
45	566	8.3	1052	1 PAK1_HUMAN	Q05397 homo sapien

ALIGNMENTS

RESULT 1	ERB2_HUMAN	STANDARD;	PRT; 1255 AA.
AC	P04626;		
AD	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)		
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell)		
DE	surface receptor HER2) (MLN 19).		
GN	ERBB2 OR HER2 OR NGL OR NEU.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID:9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86118663; PubMed=3003577;		
RA	Yamamoto T., Ikawa S., Akiyama T., Sema K., Nomura N., Miyajima N.,		
RA	Salto T., Toyoshima K.;		
RT	"Similarity of protein encoded by the human c-erb-B-2 gene to		
RT	epidermal growth factor receptor.";		
RL	Nature 319:230-234 (1986).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86070181; PubMed=2999974;		
RA	Cousens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,		
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,		
RA	Frankel U., Levinson A., Ulrich A.;		
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor		
RT	shares chromosomal location with neu oncogene.";		
RL	Science 230:1132-1139 (1985).		
RN	[3]		
RP	SEQUENCE OF 737-1031 FROM N.A.		
RX	MEDLINE=86016729; PubMed=2995967;		
RA	Sema K., Kamata N., Toyoshima K., Yamamoto T.;		
RT	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the		
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a		
RT	human salivary gland adenocarcinoma.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501 (1985).		
RN	[4]		
RP	VARIANTS VAL-654 AND VAL-655.		
RX	MEDLINE=93194196; PubMed=8095488;		
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.;		
RT	"Characterization of a new allele of the human ERBB2 gene by allele-		
RT	specific competition hybridization.";		
RL	Genomics 15:426-429 (1993).		
CC	- FUNCTION: Essential component of a neuroguin-receptor complex,		
CC	although neuroguins do not interact with it alone. GP30 is a		
CC	potential ligand for this receptor. Not activated by EGF, TGF-		
CC	alpha and amphiregulin.		
CC	- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein		
CC	tyrosine phosphate.		
CC	- SUBUNIT: Heterodimer with each of the other ERBB receptors		
CC	(Potential). Interacts with PRKCAP (By similarity).		
CC	- SUBCELLULAR LOCATION: Type I membrane protein.		

CC	-1	PMB ligand-binding increase phosphorylation on tyrosine residues (By similarity).
CC	-1	POLYMORPHISM; There are four alleles due to the variations in positions 654 and 655. Allele B1 (654-Ile-Ile-655) has a frequency of 0.782; allele B2 (654-Ile-Val-655) has a frequency of 0.206; allele B3 (654-Val-Val-655) has a frequency of 0.012.
CC	-1	SIMILARITY: BELONGS TO THE BGF RECEPTOR FAMILY.
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.ebi.ac.uk/ebis/sequence/embldb/embldb.html or send an email to license@ebi.ac.uk).
CC		-----
DR	EMBL	M11767, AAA35808.1, -
DR	EMBL	M11761, AAA35808.1, JOINED.
DR	EMBL	M11762, AAA35808.1, JOINED.
DR	EMBL	M11763, AAA35808.1, JOINED.
DR	EMBL	M11764, AAA35808.1, JOINED.
DR	EMBL	M11765, AAA35808.1, JOINED.
DR	EMBL	M11766, AAA35808.1, JOINED.
DR	EMBL	M11730, AAA75493.1, -
DR	EMBL	M12036, AAA58978.1, -
DR	EMBL	X03363, CAA27060.1, -
DR	PIR	A24571, A24571.
DR	PDB	1N8Z, 18-FEB-03.
DR	PDB	1QR1, 01-JAN-00.
DR	GeneID	HGNC:3430, ERBB2.
DR	MM	164870, -
DR	GO	GO:0005012, P:Neu/ErbB-2 receptor activity; TAS.
DR	GO	GO:0004716, P:receptor signaling protein tyrosine kinase; TAS.
DR	GO	GO:0008283, P:cell proliferation; TAS.
DR	GO	GO:0007048, P:oncogenesis; TAS.
DR	GO	GO:0006470, P:protein amino acid dephosphorylation; TAS.
DR	GO	GO:0006468, P:protein amino acid phosphorylation; TAS.
DR	InterPro	IPRO00494, EGFR_L_domain.
DR	InterPro	IPRO06211, Furin-like.
DR	InterPro	IPRO06212, Furin repeat.
DR	InterPro	IPRO00719, Prot_kinase.
DR	InterPro	IPRO01245, Tyr_Dkinase.
DR	InterPro	IPRO04019, YFP_motif.
DR	Pfam	PF00757, Furin-like; 1.
DR	Pfam	PF00069, pkinase; 1.
DR	Pfam	PF01030, Recep_L_domain; 2.
DR	Pfam	PF02757, YLP; 2.
DR	PRINTS	PR00109, TYRKINASE.
DR	PRODOM	PD000001, Prot_kinase; 1.
DR	SMART	SMO0261, FUK; 4.
DR	SMART	SMO0219, TYRC; 1.
DR	PROSITE	PS00107, PROTEIN_KINASE_ATP; 1.
DR	PROSITE	PS00109, PROTEIN_KINASE_TYR; 1.
DR	PROSITE	PS50011, PROTEIN_KINASE_DOM; 1.
KW	Transmembrane	Glycoprotein; Multigene family; Receptor; Signal;
KW	Transmembrane	Tyrosine-Protein kinase; ATP-binding; Phosphorylation;
KW	Polymorphism	3D-structure.
FT	FT SIGNAL	1..21 POTENTIAL.
FT	FT CHAIN	22..1255 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
FT	FT DOMAIN	22..652 EXTRACELLULAR (POTENTIAL).
FT	FT TRANSMEM	653..675 POTENTIAL.
FT	FT DOMAIN	676..1255 CYTOPLASMIC (POTENTIAL).
FT	FT NP_BIND	726..734 ATP (BY SIMILARITY).
FT	FT BINDING	753..753 ATP (BY SIMILARITY).
FT	FT ACT_SITE	945..945 BY SIMILARITY.
FT	FT DISULFID	195..204 BY SIMILARITY.
FT	FT DISULFID	199..212 BY SIMILARITY.
FT	FT DISULFID	220..227 BY SIMILARITY.
FT	FT DISULFID	224..235 BY SIMILARITY.
FT	FT DISULFID	236..244 BY SIMILARITY.
FT	FT DISULFID	240..252 BY SIMILARITY.
FT	FT DISULFID	255..264 BY SIMILARITY.

FT	DISULFID	268	295	311	BY SIMILARITY.
FT <td>DISULFID</td> <td>299</td> <td>311</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	299	311		BY SIMILARITY.
FT <td>DISULFID</td> <td>315</td> <td>331</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	315	331		BY SIMILARITY.
FT <td>DISULFID</td> <td>334</td> <td>338</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	334	338		BY SIMILARITY.
FT <td>DISULFID</td> <td>511</td> <td>520</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	511	520		BY SIMILARITY.
FT <td>DISULFID</td> <td>515</td> <td>528</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	515	528		BY SIMILARITY.
FT <td>DISULFID</td> <td>531</td> <td>540</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	531	540		BY SIMILARITY.
FT <td>DISULFID</td> <td>544</td> <td>560</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	544	560		BY SIMILARITY.
FT <td>DISULFID</td> <td>563</td> <td>576</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	563	576		BY SIMILARITY.
FT <td>DISULFID</td> <td>567</td> <td>584</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	567	584		BY SIMILARITY.
FT <td>DISULFID</td> <td>587</td> <td>596</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	587	596		BY SIMILARITY.
FT <td>DISULFID</td> <td>600</td> <td>623</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	600	623		BY SIMILARITY.
FT <td>DISULFID</td> <td>626</td> <td>634</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	626	634		BY SIMILARITY.
FT <td>DISULFID</td> <td>630</td> <td>642</td> <td></td> <td>BY SIMILARITY.</td>	DISULFID	630	642		BY SIMILARITY.
FT <td>MOD_RRS</td> <td>1139</td> <td>1139</td> <td></td> <td>PHOSPHORYLATION (AUTO-) (BY SIMILARITY).</td>	MOD_RRS	1139	1139		PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT <td>MOD_RRS</td> <td>1248</td> <td>1248</td> <td></td> <td>PHOSPHORYLATION (AUTO-) (BY SIMILARITY).</td>	MOD_RRS	1248	1248		PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT <td>CARBOHYD</td> <td>68</td> <td>68</td> <td></td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	68	68		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>124</td> <td>124</td> <td></td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	124	124		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>187</td> <td>187</td> <td></td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	187	187		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>259</td> <td>259</td> <td></td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	259	259		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>530</td> <td>530</td> <td></td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	530	530		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>571</td> <td>571</td> <td></td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	571	571		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>629</td> <td>629</td> <td></td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	629	629		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>VARIANT</td> <td>654</td> <td>654</td> <td></td> <td>I -> V (IN DBSNP:1801201).</td>	VARIANT	654	654		I -> V (IN DBSNP:1801201).
FT <td>VARIANT</td> <td>655</td> <td>655</td> <td></td> <td>/FTID=VAR_004077.</td>	VARIANT	655	655		/FTID=VAR_004077.
FT <td>VARIANT</td> <td>655</td> <td>655</td> <td></td> <td>I -> V (IN DBSNP:1801200).</td>	VARIANT	655	655		I -> V (IN DBSNP:1801200).
FT <td>CONFLICT</td> <td>1170</td> <td>1170</td> <td></td> <td>/FTID=VAR_004078.</td>	CONFLICT	1170	1170		/FTID=VAR_004078.
FT <td>SEQUENCE</td> <td>1255 AA;</td> <td>137909 MW;</td> <td>3995DEPDAD4DCE962 CRC64;</td> <td>P -> A (IN REF. 2).</td>	SEQUENCE	1255 AA;	137909 MW;	3995DEPDAD4DCE962 CRC64;	P -> A (IN REF. 2).
Query Match	Beet Local Similarity	100.0%;	Score 6815;	DB 1;	Length 1255;
Matches 1255;	Conservative	0;	Mismatches	0;	Indels
OY	1	MEALALCRMGILLALALPPGASASTOVCGTGMKLRPASPETHLDMRLHYOGCVOVQGNL	60		
DB	1	MEALALCRMGILLALALPPGASASTOVCGTGMKLRPASPETHLDMRLHYOGCVOVQGNL	60		
OY	61	ELTYLPTNASLSPLQDIOEVQGYVLAHNOVQVPLQRLRIVRGTQLFEDNYALAVLDNG	120		
DB	61	ELTYLPTNASLSPLQDIOEVQGYVLAHNOVQVPLQRLRIVRGTQLFEDNYALAVLDNG	120		
OY	121	DELANTTPVNGASRGGLRELQRLSTLILKGVLIQNRPOLCYQDITLMKQIFHNQOLA	180		
DB	121	DELANTTPVNGASRGGLRELQRLSTLILKGVLIQNRPOLCYQDITLMKQIFHNQOLA	180		
OY	181	LTLIDTRSRACHPCSPMGKSGRMCSSSEDCQSLRTVCGAGCARCKGRLPTDCHEQC	240		
DB	181	LTLIDTRSRACHPCSPMGKSGRMCSSSEDCQSLRTVCGAGCARCKGRLPTDCHEQC	240		
OY	241	AMGCTGKXHSDCIALCFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP	300		
DB	241	AMGCTGKXHSDCIALCFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP	300		
OY	301	YNYLSTVGSCTLVCPRLHNOEVTABDGTQCEKSKPCARVCYGLGMEHLREVAALVTSAN	360		
DB	301	YNYLSTVGSCTLVCPRLHNOEVTABDGTQCEKSKPCARVCYGLGMEHLREVAALVTSAN	360		
OY	361	IQEPAGCKKIFGSLAPLPESPDGDPASNTAPLOEPOLOVFETLLEITLYLISAMPDSL	420		
DB	361	IQEPAGCKKIFGSLAPLPESPDGDPASNTAPLOEPOLOVFETLLEITLYLISAMPDSL	420		
OY	421	DLVSFQNLQVIRGRILHNGAVSLTLQIGISWLGRLSRLRELGSGIALIHNTHLCTVHTV	480		
DB	421	DLVSFQNLQVIRGRILHNGAVSLTLQIGISWLGRLSRLRELGSGIALIHNTHLCTVHTV	480		
OY	481	PMQDLFRNPHOALAHANRPEDECVGEGALACHQICARGHCMGPRPTQCVNCSQFLRQEC	540		
DB	481	PMQDLFRNPHOALAHANRPEDECVGEGALACHQICARGHCMGPRPTQCVNCSQFLRQEC	540		
OY	541	VVEKRVYQGLPREVYNARHCLPCGRPEOPONGSYTCFPRBADOCVACHYADPPCVARC	600		
DB	541	VVEKRVYQGLPREVYNARHCLPCGRPEOPONGSYTCFPRBADOCVACHYADPPCVARC	600		

QY 601 PSQVDPDLSYMPKWPDEBEGACQPCPINCTHSCVDLDDKCGPAQORASPLTISIAVVG 660
 DB 601 PSQVDPDLSYMPKWPDEBEGACQPCPINCTHSCVDLDDKCGPAQORASPLTISIAVVG 660
 QY 661 ILLVVVGVVFGIILIKRQOKIRKTYMRLLQETLVEPLTSSGAMPQAOQRILIKETEL 720
 DB 661 ILLVVVGVVFGIILIKRQOKIRKTYMRLLQETLVEPLTSSGAMPQAOQRILIKETEL 720
 QY 721 RKVKVLSGAGFVYKGIWIPDGENVKIPVAIKVLRENTSPKANBEILDEAVYMGVSGP 780
 DB 721 RKVKVLSGAGFVYKGIWIPDGENVKIPVAIKVLRENTSPKANBEILDEAVYMGVSGP 780
 QY 781 YVSRLLGICLTSTVOLVTLMPYGGCLLDHVRNRRGLSGQDLNMCQIACKMSYLEDVR 840
 DB 781 YVSRLLGICLTSTVOLVTLMPYGGCLLDHVRNRRGLSGQDLNMCQIACKMSYLEDVR 840
 QY 841 LVHDDLAAARNLVKSPNNVKTIDBGLARLLDDETEHYADGGKVPKMMALSLIKRRFT 900
 DB 841 LVHDDLAAARNLVKSPNNVKTIDBGLARLLDDETEHYADGGKVPKMMALSLIKRRFT 900
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWVCKM 960
 DB 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWVCKM 960
 QY 961 IDSECRPRFRLVSEFSHWARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDMDGLDVA 1020
 DB 961 IDSECRPRFRLVSEFSHWARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDMDGLDVA 1020
 QY 1021 BEYLVPOQGFPCPDPAFAGGWHHRHSSSTRSGGDLTLGLBSEBAPRSLAPSEBG 1080
 DB 1021 BEYLVPOQGFPCPDPAFAGGWHHRHSSSTRSGGDLTLGLBSEBAPRSLAPSEBG 1080
 QY 1081 AGSDVFDGDLGGAAGKGLQSLPTHDPSPLOKYSBPPTVPLPSETGYVAPLTCSPQPEYV 1140
 DB 1081 AGSDVFDGDLGGAAGKGLQSLPTHDPSPLOKYSBPPTVPLPSETGYVAPLTCSPQPEYV 1140
 QY 1141 NQPDVPRPQPSPRREGPLPAARPAATLRLPKTLSPKNGVVDVAFGAVENPEYLTPO 1200
 DB 1141 NQPDVPRPQPSPRREGPLPAARPAATLRLPKTLSPKNGVVDVAFGAVENPEYLTPO 1200
 QY 1201 CGAAPQHPHPPAFSPFNLVWDDPPBRKAPSESTFKTPTAENPEYLTGLDVPV 1255
 DB 1201 CGAAPQHPHPPAFSPFNLVWDDPPBRKAPSESTFKTPTAENPEYLTGLDVPV 1255

RESULT 2
 ID ERB2_RAT STANDARD; PRT; 1257 AA.
 AC P06494;
 DT 01-JAN-1998 (Rel. 06, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor receptor-related protein).
 DE ERBB2 OR NEU.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Neuroblastoma;
 RX MEDLINE=86118662; PubMed=3945311;
 RA Bergmann C.I., Hung M.-C., Weinberg R.A.;
 RT "The neu oncogene encodes an epidermal growth factor receptor-related protein."
 RL Nature 319:226-230(1986).
 RP [2]
 RP SEQUENCE OF 852-905 FROM N.A.
 RC TISSUE=Sciatic nerve;
 RX MEDLINE=91222560; PubMed=2025425;

RA Lat C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially
 RT expressed in the vertebrate nervous system";
 RL Neuron 6:691-704(1991).
 RN [3]
 RP STRUCTURE BY NMR OF 650-668.
 RX MEDLINE=92155181; PubMed=1346763;
 RA Gullick W.J., Bottomley A.C., Lofes F.J., Doak D.G., Mulvey D.,
 RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 RT "Three dimensional structure of the transmembrane region of the proto-
 RT oncogenic and oncogenic forms of the neu protein";
 RL EMBO J. 11:443-448(1992).
 CC -1- FUNCTION: Essential component of a neurogulin-receptor complex,
 CC although neurogulins do not interact with it alone. GP30 is a
 CC potential ligand for this receptor. Not activated by EGF, TGF-
 CC alpha and amphiregulin.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: Heterodimer with each of the other ERBB receptors. The
 CC constitutively activated oncogenic variant forms a homodimer.
 CC Interacts with PRKCAP (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine
 CC residues (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, X03362; CAA27059.1; ALT_INIT.
 CC PDB: 1ITJ; 27-JUN-01.
 DR PDB: 1NB7; 18-FEB-03.
 DR InterPro: IPR000494; EGFR_L domain.
 DR InterPro: IPR006211; Furin-like.
 DR InterPro: IPR006212; Furin repeat.
 DR InterPro: IPR000719; Prot. Kinase.
 DR InterPro: IPR001245; Tyr. Kinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF00069; Pkinase; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR Pfam: PF02757; YLP; 2.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_Kinase; 1.
 DR SMART: SM00261; FU; 4.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Proto-oncogene; Disease mutation; 3D-structure.
 FT CHAIN 1..21
 FT SIGNAL 1..21
 FT DOMAIN 22..1257 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 FT TRANSMEM 655..677 POTENTIAL.
 FT DOMAIN 678..1257 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 159..369 CYS-RICH.
 FT DOMAIN 473..646 CYS-RICH.
 FT DOMAIN 722..989 PROTEIN KINASE.
 FT NP_BIND 728..736 ATP (BY SIMILARITY).
 FT BINDING 755..755 ATP (BY SIMILARITY).
 FT ACT_SITE 847..847 BY SIMILARITY.
 FT DISULFID 196..205 BY SIMILARITY.
 FT DISULFID 200..213 BY SIMILARITY.
 FT DISULFID 221..228 BY SIMILARITY.
 FT DISULFID 225..236 BY SIMILARITY.
 FT DISULFID 237..245 BY SIMILARITY.

FT DISULFID 241 253 BY SIMILARITY.
 FT DISULFID 256 265 BY SIMILARITY.
 FT DISULFID 269 296 BY SIMILARITY.
 FT DISULFID 300 312 BY SIMILARITY.
 FT DISULFID 316 332 BY SIMILARITY.
 FT DISULFID 335 339 BY SIMILARITY.
 FT DISULFID 513 522 BY SIMILARITY.
 FT DISULFID 517 530 BY SIMILARITY.
 FT DISULFID 533 542 BY SIMILARITY.
 FT DISULFID 546 562 BY SIMILARITY.
 FT DISULFID 565 578 BY SIMILARITY.
 FT DISULFID 569 586 BY SIMILARITY.
 FT DISULFID 589 598 BY SIMILARITY.
 FT DISULFID 602 625 BY SIMILARITY.
 FT DISULFID 628 636 BY SIMILARITY.
 FT DISULFID 632 644 BY SIMILARITY.
 FT MOD_RES 1141 1141 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1250 1250 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).
 SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;

Query Match 88.1%; Score 6003; DB 1; Length 1257;
 Best Local Similarity 88.0%; Pred. No. 7, 6e-310;

Matches 1106; Conservative 49; Mismatches 100; Indels 2; Gaps 2;

QY 1 MELALCRMGLTALLPFGASTOVCTGTDMRLPASPETHLDMRLHLYGCCQVVGNTL 60
 DB 1 MELAMCMGFLTALLPFGIAGTQCTGTDMRLPASPETHLDMRLHLYGCCQVVGNTL 60
 QY 61 ELTTLPTAASISFLDIOEVQSVTIANQVQVPLQRLIRYRGTQLFEDNTALAVLNG 120
 DB 61 ELTYPAVASISFLDIOEVQSVTIANQVQVPLQRLIRYRGTQLFEDNTALAVLNR 120
 QY 121 DPLNNTTFTV- GASPGGLREIQLRSITELIKGGLVLIORPOLCYODTIMKDI FHKNNOL 179
 DB 121 DPODVAASTFRTYRGLREIQLRSITELIKGGLVLIORPOLCYODTIMKDI FHKNNOL 180
 QY 180 ALLTIDNRSRACHP CSPMCKSGRCMGSSEDCOSLTRITVCAAGCARCKGPLPTDCBHQ 239
 DB 181 APVDIDNRSRACHP CAPCKONHMGSSPEBCOILGITICSGCARCKGRLPTDCBHQ 240
 QY 240 CAAGTGRHSDCLALCPHNSGICELHCPALVTNTDTFESMPNREGRTYFGASCVTAC 299
 DB 241 CAAGTGRHSDCLALCPHNSGICELHCPALVTNTDTFESMPNREGRTYFGASCVTTC 300
 QY 300 PNYLSTVGSGCTIACPLHNOEVTABDGTORCEKSKPCARCYGLGMEHLREVAAYSA 359
 DB 301 PNYLSTVGSGCTIACPLHNOEVTABDGTORCEKSKPCARCYGLGMEHLREVAAYSD 360
 QY 360 NIOBAGCKIIFGSLAFLPESFGDPASNTAFLQEPQLQVFTLEITGYLYISAMPDYL 419
 DB 361 NVQBFPGCKIIFGSLAFLPESFGDPSSGIALPLRPEQLQVFTLEITGYLYISAMPDYL 420
 QY 420 PDLSTFQNLQVIRKGIILHNGAVSLTLOGLISWLGIRLSIRELSGLALIHNTHLCPVHT 479
 DB 421 RDLSTFQNLQVIRKGIILHNGAVSLTLOGLISWLGIRLSIRELSGLALIHNTHLCPVHT 480
 QY 480 VPMDDLFRNPHOALHTANRPEDE- CVBGLGCHOLCAHGCHMGPGPTCCVACSGPRLRQ 538
 DB 481 VPMDDLFRNPHOALHTANRPEDE- CVBGLGCHOLCAHGCHMGPGPTCCVACSGPRLRQ 540
 QY 539 ECVEBCEVLOGLPREYVNAHCLCPHPECOPONGSVTCGPBAPDQVACAHYKDDPFCVA 598
 DB 541 ECVEBCEVLOGLPREYVNAHCLCPHPECOPONGSVTCGPBAPDQVACAHYKDDPFCVA 600
 QY 599 RCPGSGVPLSLMPTWKPPDEBACOPPCINCTHSCVDLDDYGCAPBAPASPLTSLIISAV 658

DB 601 RCPGSGVPLSLMPTWKPPDEBACOPPCINCTHSCVDLDDYGCAPBAPASPLTSLIATV 660
 QY 659 VGIILLVVLGVVFGILLIKRQOKIRKYMRLLOTELVEPLTPSGAMPNOAMILKET 718
 DB 661 VGVLLFLLVVVGVVFGILLIKRQOKIRKYMRLLOTELVEPLTPSGAMPNOAMILKET 720
 QY 719 ELRKRYLGSAGFTVYKGIWIPGSENVKIPVALKYLBENTSPKANKELIDAYVMAGVG 778
 DB 721 ELRKRYLGSAGFTVYKGIWIPGSENVKIPVALKYLBENTSPKANKELIDAYVMAGVG 780
 QY 779 SPYVSRILGICLTSTVOLVTOIMPYGLLDHYRENRGLSGODLNMCMQIAKMSYLED 838
 DB 781 SPYVSRILGICLTSTVOLVTOIMPYGLLDHYRENRGLSGODLNMCMQIAKMSYLED 840
 QY 839 VRLVHRDLAARNVLYKSNHVKITDPGLARLLDIDETEYHADGCKVPIKMMALLESILRR 898
 DB 841 VRLVHRDLAARNVLYKSNHVKITDPGLARLLDIDETEYHADGCKVPIKMMALLESILRR 900
 QY 899 FTHOSDWSYGVTTWELMTFGAKPYDGIIPAREIPDLKGERLPORPCTIDYVIMIVKC 958
 DB 901 FTHOSDWSYGVTTWELMTFGAKPYDGIIPAREIPDLKGERLPORPCTIDYVIMIVKC 960
 QY 959 WMISECPRRRELVSERSRMAADPQRFVIONEDLGASPLDSTFFYSLEDDMDGLV 1018
 DB 961 WMISECPRRRELVSERSRMAADPQRFVIONEDLGASPLDSTFFYSLEDDMDGLV 1020
 QY 1019 DAEELVYPOGFCFDPAPAGAGVHHRSSSTRSGGDLTLGLPSEBEAPRSLAPS 1078
 DB 1021 DAEELVYPOGFCFDPAPAGAGVHHRSSSTRSGGDLTLGLPSEBEAPRSLAPS 1080
 QY 1079 EGAGSDVFDGLMGCAKGLQSLPTHDPSPLOKSEDETPVLPSTDTGVAPLTPSGPOE 1138
 DB 1081 EGAGSDVFDGLMGCAKGLQSLPTHDPSPLOKSEDETPVLPSTDTGVAPLTPSGPOE 1140
 QY 1139 YVNOEDVPOPPSPREGPLPAPAPAGATLERPKTSPKNGVYKVPAPFGAVENPEVLT 1198
 DB 1141 YVNOEDVPOPPSPREGPLPAPAPAGATLERPKTSPKNGVYKVPAPFGAVENPEVLT 1200
 QY 1199 PGGGAAPQHPHPAPFSPAFDNLVYWDOPPERGAPPTSTFKGTPTANBEVYLGLDVPV 1255
 DB 1201 PREGTASPHSPAPFSPAFDNLVYWDQSSSQGPPSPFEGTPTANBEVYLGLDVPV 1257

RESULT 3
 ERB2_MESAU
 ID ERB2_MESAU STANDARD; PRT; 1254 AA.
 AC Q60553;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
 ID (p185erbB2) (NEU proto-oncogene) (C-erbB-2).
 GN ERB2 OR NEU.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 CC Mesocricetus.
 OX NCBI_TaxID:10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Nerve;
 RX MEDLINE=94193007; PubMed=7908275;
 RA Nakamura Y, Ushijima T, Ishizaka Y, Nagao M, Arai M,
 RA Yamazaki Y, Ishikawa T;
 RT "Cloning and activation of the Syrian hamster neu proto-oncogene."
 RL Gene 140:251-255(1994).
 CC -I- FUNCTION: Essential component of a neurogulin-receptor complex,
 CC although neurogulin does not interact with it alone. GP30 is a
 CC potential ligand for this receptor. Not activated by EGF, TGF-
 CC alpha and amphiregulin (By similarity).
 CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -I- SUBUNIT: Heterodimer with each of the other ERB2 receptors

CC (Potential). Interacts with PRKCAP (By similarity).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine residues.

CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.embnet.org/licenses/embnet.html>).

CC or send an email to license@ebi.ac.uk.

CC -----

DR EMBL; D16295; BAA03801.1; -.

DR PIR; I48161; I48161.

DR HSPG; P1362; 1FGK.

DR InterPro; IPR000494; EGFR_L_domain.

DR InterPro; IPR006211; Furin-Like.

DR InterPro; IPR006212; Furin_repeat.

DR InterPro; IPR000719; Prot_Kinase.

DR InterPro; IPR001245; Tyr_Kinase.

DR InterPro; IPR004019; YLP_motif.

DR Pfam; PF00757; Furin_Like; 1.

DR Pfam; PF00069; Kinase; 1.

DR Pfam; PF01030; Recep_L_domain; 2.

DR Pfam; PF02757; YLP; 2.

DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Prot_Kinase; 1.

DR SMART; SMO0261; FU; 4.

DR SMART; SMO0219; Tyrc; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.

DR Transmembrane; Glycoprotein; Multigene family; Receptor; Signal; Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation; K4 proto-oncogene; Disease mutation.

K4 SIGNAL 1 21

FT CHAIN 22 1254

FT DOMAIN 22 652

FT TRANSMEM 653 675

FT DOMAIN 676 1254

FT DOMAIN 158 368

FT DOMAIN 472 644

FT DOMAIN 720 987

FT NP_BIND 726 734

FT BINDING 753 753

FT ACT_SITE 845 845

FT DISULFID 195 204

FT DISULFID 199 212

FT DISULFID 236 244

FT DISULFID 240 252

FT DISULFID 255 264

FT DISULFID 268 295

FT DISULFID 299 311

FT DISULFID 315 331

FT DISULFID 334 338

FT DISULFID 511 520

FT DISULFID 515 528

FT DISULFID 531 540

FT DISULFID 544 560

FT DISULFID 563 576

FT DISULFID 567 584

FT DISULFID 587 596

FT DISULFID 600 623

FT DISULFID 626 634

FT DISULFID 630 642

FT MOD_RES 1139 1139

FT MOD_RES 1247 1247

FT CARBOHYD 68 68

FT CARBOHYD 125 125

FT CARBOHYD 187 187

FT CARBOHYD 259 259

FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARIANT 658 658 V -> E (IN ONCOGENIC NEU).

FT VARIANT 659 659 V -> E (IN ONCOGENIC NEU).

FT SEQUENCE 1254 AA; 138252 MW; 974C3791C21F2BE1 CRC64;

Query Match 87.9%; Score 5993.5; DB 1; Length 1254;

Best Local Similarity 87.7%; Pred. No. 2.4e-309;

Matches 1101; Conservative 57; Mismatches 96; Indels 1; Gaps 1;

1 MELALCWGILLALLPFGAASVQVCTGTMDLRLPASBETHLDIVRHLYGCGVQVGNL 60

1 MELAMCGWGLLLALSPGASGTQVCTGTMDLRLPASBETHLDIVRHLYGCGVQVGNL 60

61 ELTYLPTVASLPLQDIEVQGVYVLIANQVQVQLRRLIRVGTQLFEDNALAVLDNG 120

61 ELTYLPTVASLPLQDIEVQGVYVLIANQVQVQLRRLIRVGTQLFEDNALAVLDNG 120

61 ELTYLPTVASLPLQDIEVQGVYVLIANQVQVQLRRLIRVGTQLFEDNALAVLDNG 120

121 PLNNTPVTGASPGGLREQLRLSLTEILKGVLIQRNPQLCYODTILMKDIPKQNOQLA 180

121 PLNNTPVTGASPGGLREQLRLSLTEILKGVLIQRNPQLCYODTILMKDIPKQNOQLA 180

121 PLNNTPVTGASPGGLREQLRLSLTEILKGVLIQRNPQLCYODTILMKDIPKQNOQLA 180

181 LTLIDTNSRACPSCSPKCKSGRCWGESSEDCQSLTRYTCAGGACRCCKPLPTDCHEQC 240

181 LTLIDTNSRACPSCSPKCKSGRCWGESSEDCQSLTRYTCAGGACRCCKPLPTDCHEQC 240

181 LTLIDTNSRACPSCSPKCKSGRCWGESSEDCQSLTRYTCAGGACRCCKPLPTDCHEQC 240

241 AAGCTGPGHSDCLCFHNSGICELHCPALVYNTDTFESMPNEGRTYFASCVTACP 300

241 AAGCTGPGHSDCLCFHNSGICELHCPALVYNTDTFESMPNEGRTYFASCVTACP 300

241 AAGCTGPGHSDCLCFHNSGICELHCPALVYNTDTFESMPNEGRTYFASCVTACP 300

301 YNYLSTVGSCTLVCPPLNQBVTADGTQRCCKSKPCARVCYGLGMEHLREAVATSAN 360

301 YNYLSTVGSCTLVCPPLNQBVTADGTQRCCKSKPCARVCYGLGMEHLREAVATSAN 360

301 YNYLSTVGSCTLVCPPLNQBVTADGTQRCCKSKPCARVCYGLGMEHLREAVATSAN 360

301 YNYLSTVGSCTLVCPPLNQBVTADGTQRCCKSKPCARVCYGLGMEHLREAVATSAN 360

301 YNYLSTVGSCTLVCPPLNQBVTADGTQRCCKSKPCARVCYGLGMEHLREAVATSAN 360

361 IQEPAGCKKIGSLAFLESFSDGNSGSIAPLTPBOLQFETLEITGLYISAWPDSLSH 420

361 IQEPAGCKKIGSLAFLESFSDGNSGSIAPLTPBOLQFETLEITGLYISAWPDSLSH 420

361 IQEPAGCKKIGSLAFLESFSDGNSGSIAPLTPBOLQFETLEITGLYISAWPDSLSH 420

421 DLSVQNLQVIRGRILHNGAVSLTQGLISWLGRLSRLSRLSSGLALIHNTHLCEVHTV 480

421 DLSVQNLQVIRGRILHNGAVSLTQGLISWLGRLSRLSRLSSGLALIHNTHLCEVHTV 480

421 DLSVQNLQVIRGRILHNGAVSLTQGLISWLGRLSRLSRLSSGLALIHNTHLCEVHTV 480

421 DLSVQNLQVIRGRILHNGAVSLTQGLISWLGRLSRLSRLSSGLALIHNTHLCEVHTV 480

421 DLSVQNLQVIRGRILHNGAVSLTQGLISWLGRLSRLSRLSSGLALIHNTHLCEVHTV 480

481 PMDQFRPHQALHTANRPEDECVGEGALCHQLCARGHCHWGPPTQVNCQSFRLGQSC 540

481 PMDQFRPHQALHTANRPEDECVGEGALCHQLCARGHCHWGPPTQVNCQSFRLGQSC 540

481 PMDQFRPHQALHTANRPEDECVGEGALCHQLCARGHCHWGPPTQVNCQSFRLGQSC 540

481 PMDQFRPHQALHTANRPEDECVGEGALCHQLCARGHCHWGPPTQVNCQSFRLGQSC 540

481 PMDQFRPHQALHTANRPEDECVGEGALCHQLCARGHCHWGPPTQVNCQSFRLGQSC 540

541 VECRCVLOGLPREYVNAHCLPCHRECPQNSVTCFGBADQCVACAHYKPPFCVANC 600

541 VECRCVLOGLPREYVNAHCLPCHRECPQNSVTCFGBADQCVACAHYKPPFCVANC 600

541 VECRCVLOGLPREYVNAHCLPCHRECPQNSVTCFGBADQCVACAHYKPPFCVANC 600

541 VECRCVLOGLPREYVNAHCLPCHRECPQNSVTCFGBADQCVACAHYKPPFCVANC 600

541 VECRCVLOGLPREYVNAHCLPCHRECPQNSVTCFGBADQCVACAHYKPPFCVANC 600

601 PSQVSPDSTYPIKPFEBEGACOPCPINCHSCVDLDKGPABORASPLTSISAVVG 660

601 PSQVSPDSTYPIKPFEBEGACOPCPINCHSCVDLDKGPABORASPLTSISAVVG 660

601 PSQVSPDSTYPIKPFEBEGACOPCPINCHSCVDLDKGPABORASPLTSISAVVG 660

601 PSQVSPDSTYPIKPFEBEGACOPCPINCHSCVDLDKGPABORASPLTSISAVVG 660

601 PSQVSPDSTYPIKPFEBEGACOPCPINCHSCVDLDKGPABORASPLTSISAVVG 660

661 ILVVLVGVVGLIKRQOKIKRYTMRRLQETELVBLTPSGAMPNQAKRILKETEL 720

661 ILVVLVGVVGLIKRQOKIKRYTMRRLQETELVBLTPSGAMPNQAKRILKETEL 720

661 ILVVLVGVVGLIKRQOKIKRYTMRRLQETELVBLTPSGAMPNQAKRILKETEL 720

661 ILVVLVGVVGLIKRQOKIKRYTMRRLQETELVBLTPSGAMPNQAKRILKETEL 720

661 ILVVLVGVVGLIKRQOKIKRYTMRRLQETELVBLTPSGAMPNQAKRILKETEL 720

721 RKVYVLSGAGTGYKGIWIPDEGENVKIPLAIVLRENTSPRANKELIDEAVYAGVSP 780

721 RKVYVLSGAGTGYKGIWIPDEGENVKIPLAIVLRENTSPRANKELIDEAVYAGVSP 780

721 RKVYVLSGAGTGYKGIWIPDEGENVKIPLAIVLRENTSPRANKELIDEAVYAGVSP 780

721 RKVYVLSGAGTGYKGIWIPDEGENVKIPLAIVLRENTSPRANKELIDEAVYAGVSP 780

721 RKVYVLSGAGTGYKGIWIPDEGENVKIPLAIVLRENTSPRANKELIDEAVYAGVSP 780

781 YVSRLLGICLSTVQVLTQVLMFYGLLHVRENRRGLSODLNNCOMQAKAGSYLEDVR 840

781 YVSRLLGICLSTVQVLTQVLMFYGLLHVRENRRGLSODLNNCOMQAKAGSYLEDVR 840

781 YVSRLLGICLSTVQVLTQVLMFYGLLHVRENRRGLSODLNNCOMQAKAGSYLEDVR 840

781 YVSRLLGICLSTVQVLTQVLMFYGLLHVRENRRGLSODLNNCOMQAKAGSYLEDVR 840

781 YVSRLLGICLSTVQVLTQVLMFYGLLHVRENRRGLSODLNNCOMQAKAGSYLEDVR 840

841 LVHRDLAARNVLYSPNHKITTDFGLARLLDIDETEHADGKVPKIMNALSILRRRT 900

841 LVHRDLAARNVLYSPNHKITTDFGLARLLDIDETEHADGKVPKIMNALSILRRRT 900

841 LVHRDLAARNVLYSPNHKITTDFGLARLLDIDETEHADGKVPKIMNALSILRRRT 900

841 LVHRDLAARNVLYSPNHKITTDFGLARLLDIDETEHADGKVPKIMNALSILRRRT 900

841 LVHRDLAARNVLYSPNHKITTDFGLARLLDIDETEHADGKVPKIMNALSILRRRT 900

901 HOSDWSYGVTVVLMATFGAKPYDGIPLAREIPDLLKSGRLPOPICTIDVYIMYKCMW 960

901 HOSDWSYGVTVVLMATFGAKPYDGIPLAREIPDLLKSGRLPOPICTIDVYIMYKCMW 960

901 HOSDWSYGVTVVLMATFGAKPYDGIPLAREIPDLLKSGRLPOPICTIDVYIMYKCMW 960

901 HOSDWSYGVTVVLMATFGAKPYDGIPLAREIPDLLKSGRLPOPICTIDVYIMYKCMW 960

901 HOSDWSYGVTVVLMATFGAKPYDGIPLAREIPDLLKSGRLPOPICTIDVYIMYKCMW 960

Db 901 HOSDWSYGVTVWIMLTFGANKPYDGIIPARELPDLLEKGERLPQPICTIDVIMNKCMM 960

Qy 961 IDSCRFPRFRRLVSEFSHMAADPQRFVYIQNEDLGPASPLDSTFRSLLEDMDMDLVDA 1020

Db 961 IDSCRFPRFRRLVSEFSHMAADPQRFVYIQNEDLGPASPLDSTFRSLLEDMDMDLVDA 1020

Qy 1021 EBYLVPPQGFPCPDPAAGCVGHHRRSSSTRSGGSDLTGLLEPSEEDAPRSLPASEG 1080

Db 1021 EBYLVPPQGFPCPDPAAGCVGHHRRSSSTRSGGSDLTGLLEPSEEDAPRSLPASEG 1080

Qy 1081 AGSDVFPDLDLGMGAAGKQSLPETHDPSPLQRYSEBPTVPLPSETGCVYAPLTCSPQERYV 1140

Db 1081 AGSDVFPDLDLGMGAAGKQSLPETHDPSPLQRYSEBPTVPLPSETGCVYAPLTCSPQERYV 1140

Qy 1141 NOPDVPRPQPSPREGGLPAARPAAGATLSPKTLSPGKGVVQVPAFGAVENPEYLTPQ 1200

Db 1141 NOPDVPRPQPSPREGGLPAARPAAGATLSPKTLSPGKGVVQVPAFGAVENPEYLTPQ 1200

Qy 1201 GGAAPQPPHPPAFSPAPDNLYYMDQDPPERGAPSPSTFKGTPTAENPEYLGLDVPV 1255

Db 1201 GGAAPQPPHPPAFSPAPDNLYYMDQDPPERGAPSPSTFKGTPTAENPEYLGLDVPV 1255

RESULT 4

ID EGFR HUMAN STANDARD; PRT; 1210 AA.

AC P00533; 000688; 000732; P06268; 014225; Q92795; Q9B822; Q9GZX1;

AC Q9H2C3; Q9H3C9; Q9UMD7; Q9UMD8; Q9UMG5;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor protein-tyrosine kinase ErbB-1).

OS EGFR OR ERBB1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RX MEDLINE=84219729; PubMed=6328312;

RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W., Lee J., Varden Y., Libermann T.A., Schlessinger J., Downward J., Mayes B.L.V., Whittle N., Waterfield M.D., Seeburg P.H.; "Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells."; Nature 309:418-425 (1984).

RL [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX TISSUE=Placenta;

RX MEDLINE=95382957; PubMed=7554368;

RA Ilekis J.V., Stark B.C., Scoccia B.; "Possible role of variant RNA transcripts in the regulation of epidermal growth factor receptor expression in human placenta."; Mol. Reprod. Dev. 41:149-156 (1995).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX TISSUE=Placenta;

RX MEDLINE=97078686; PubMed=8918811;

RA Reller J.L., Maibach N.J.; "A 1.8 kb alternative transcript from the human epidermal growth factor receptor gene encodes a truncated form of the receptor."; Nucleic Acids Res. 24:4050-4056 (1996).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX TISSUE=Placenta;

RX MEDLINE=97256547; PubMed=9103388;

RA Ilekis J.V., Gariti J., Niederberger C., Scoccia B.; "Expression of a truncated epidermal growth factor receptor-like protein (TEGFR) in ovarian cancer."; Gynecol. Oncol. 65:36-41 (1997).

RL [5]

RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).

RC TISSUE=Placenta;

RX MEDLINE=21100872; PubMed=11161793;

RA Reller J.L., Thredgill D.W., Eley G.D., Strunk K.B., Danielson A.J., Schehl Sinclair C., Pearshall R.S., Green P.J., Yee D., Lampland A.L., Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D., Maibach N.J.; "Comparative genomic sequence analysis and isolation of human and mouse alternative EGFR transcripts encoding truncated receptor isoforms."; Genomics 71:1-20 (2001).

RL [6]

RP SEQUENCE OF 575-687 FROM N.A.

RA Reller J.L., Thredgill D.W., Danielson A.J., Schehl C.M., Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R., Maibach N.J.; "Human and mouse alternative EGFR transcripts encoding only the extracellular domain of the receptor."; Submitted (Feb-1999) to the EMBL/Genbank/DBJ databases.

RL [7]

RP SEQUENCE OF 713-924 FROM N.A.

RX MEDLINE=84196372; PubMed=6326261;

RA Lin C.R., Chen W.S., Knutiger W., Stolarczyk L.S., Weber W., Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.; "Expression cloning of human EGF receptor complementary DNA: gene amplification and three related messenger RNA products in A431 cells."; Science 224:843-848 (1984).

RL [8]

RP SEQUENCE OF 150-962 FROM N.A.

RX MEDLINE=84245835; PubMed=6330563;

RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P., Roe B.A., Merlino G.T., Pastan I.; "Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs overproduced in A431 carcinoma cells."; Nature 309:806-810 (1984).

RL [9]

RP SEQUENCE OF 1028-1210 FROM N.A.

RX MEDLINE=85046483; PubMed=6093780;

RA Simmen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G., O'Malley B.W.; "Isolation of an evolutionarily conserved epidermal growth factor receptor cDNA from human A431 carcinoma cells."; Biochem. Biophys. Res. Commun. 124:125-132 (1984).

RL [10]

RP SEQUENCE OF 1-29 FROM N.A.

RX MEDLINE=86217333; PubMed=3329716;

RA Haley J.D., Whittle N., Bennett P., Kitchington D., Ullrich A., Waterfield M.D.; "The human EGF receptor gene: structure of the 110 kb locus and identification of sequences regulating its transcription."; Oncogene Res. 1:375-396 (1987).

RL [11]

RP SEQUENCE OF 1-29 FROM N.A.

RX MEDLINE=91107677; PubMed=1988448;

RA Haley J.D., Waterfield M.D.; "Contributory effects of de novo transcription and premature transcript termination in the regulation of human epidermal growth factor receptor proto-oncogene RNA synthesis."; J. Biol. Chem. 266:1746-1753 (1991).

RL [12]

RP SEQUENCE OF 1-29 FROM N.A.

RX MEDLINE=85270438; PubMed=2991899;

RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.; "Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene."; Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924 (1985).

RL [13]

RP SEQUENCE OF 540.

RA Kohda D.; Submitted (SEP-1997) to the SWISS-PROT data bank.

RL [14]

RP RECEPTOR ACTIVITY;

RX MEDLINE=84191554; PubMed=6325948;

Query Match	46.5%: Score 3167; DB 1; Length 1210;
Best Local Similarity 49.8%; Pred. No. 5e-160;	
Matches 632; Conservative 176; Mismatches 346; Indels 116; Gaps 22;	
CC IsoId=P00533-4; Sequence=VSP_002891, VSP_002892;	
CC -1- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also	
CC expressed in ovarian cancers.	
CC -1- MISCELLANEOUS: Binding of EGFR to the receptor leads to	
CC dimerization, internalization of the EGFR-receptor complex,	
CC induction of the tyrosine kinase activity, stimulation of cell DNA	
QY 11 LILALPFGAA--STQCTGTDMKLRAPAPETHMLRLHYQCCQVQVQGLTLYPTN 68	
DB 14 LIALCASRALBKKKCCGTSNKLQGLGFEEHFLSLQMFNNCEVALLGLTETTYQRN 73	
QY 69 ASLSFLDDIDQVQYVLIANNQVQVLOLRATYRQQLPEDIYVALVLNDGDLNNTTP 128	
DB 74 YDLSFLKTIQVQYVLIANNQVQVLOLRATYRQQLPEDIYVALVLNDGDLNNTTP 126	
QY 129 VTGASPGGLRELQRLSLTEILKSGVLIQRNPOLCYQDTILMKDIFHKNOALATLIDTRN 188	
DB 127 ---ANKTGLKELPMRNLOETILHGAVERPNNPALCNVDSIQMRDVSDFLNSMDPNH 183	
QY 189 SRACHPSSPMCKSKSRCKGSSSDCCOSTIRYVCAAGCA-RCKGPIPTCCHEQCAAGCTGP 247	
DB 184 LGSQOKDPCSPGSCMGAGEBNCKLTKLCAQCCSGRGRKSPSCCHQCACGCTGP 243	
QY 248 KHSDDCLALCFNHSGLIETELCPALVTYNTDTPFSMPNBEGRYTFGASCTVACPYNYSTD 307	
DB 244 RESDCLVCRKFRBDATKTCPLMLYNTPTTQMDVNPBGKISFGATCVKCKPNNYVTD 303	
QY 308 VGSCTIVCPRLHNOEVTAEIDTORCEKSKCAEVCYGLGMEHLREVAVTSANTQEPAGC 367	
DB 304 HGSQVRAAGADSYEM-EDGVKRCCKCEGRCAKCNQIGIGEFADSISINATNIKFRNC 362	
QY 368 KTIPLSLAFIPESPDGPASNTKAPLOEQLQVETLEETIGYLIYSMPDLSPLSPON 427	
DB 363 TSLSGDHLIPVAFRSGSPFTHPPDLPOEIDLKTVEIRIGFLLIQAMPENRDLHAFFN 422	
QY 428 LQYIRGRILHNGVYSLTLQGLISMLGRSLRELGSGLAIHHNTHLCFHYTWDOLF 487	
DB 423 LEIRIGTKOHQGRSLAVNSLNTSLTSLGRSLKEISDDAVLISGKNLCYANNTIMWKKLFG 482	
QY 488 NHPQALHTANREDEDECVSEGLACHOLCAHGMCGPPTCVNCSQFLRGOECVCEGRVL 547	
DB 483 TSGQTKLISNRGENSKATGOVCHALCSPEGCMGPBPCVCSRNYSKREKCDKCNLL 542	
QY 548 QGLPREYVNAHCLPCHPECOPONGSVTCFEPADOCVACAHYKDPFVCARCPSGVKPD 607	
DB 543 EGEPRFEVNSECIQCHPECLPQMNITCTGRPDNDCICAHYIDGHCYKTCFAGVNGE 602	
QY 608 LSTMPITMKPFDEBGAQCPRCINCTHSCVDLDDKGCRAQASPLTSLISAVNG---LLV 664	
DB 603 NNTL-VKWDADAHVCHLCPNCTCYGTGGLGECPCPNGRKIP-STATWVGLLLLLV 659	
QY 665 VLVGVVFGILIKRQOKIRKYTKRRLLOETLEVEPLTPSGAMPNOAMRLKETELRKVK 724	
DB 660 VLVGIG---LPMRRHIVKRTLRLLQREHLYVEPLTPSEANQALLRLIKETEPKIK 716	
QY 725 VLSGAGFYVYKGIWIPDGENVXIPVALIKVLRNTSPKANKELLDEAYVNAVGSPVYSR 784	
DB 717 VLSGAGFYVYKGIWIPDGENVXIPVALIKVLRNTSPKANKELLDEAYVNAVDNPHVC 776	
QY 785 LILGICLSTIYVQLMPYGCCLDHYRNRGRGLSGDILMWCOIANGSMYLEVRLVHR 844	
DB 777 LILGICLSTIYVQLMPYGCCLDHYRNRGRGLSGDILMWCOIANGSMYLEVRLVHR 836	
QY 845 DLNARNVLYSPNHYKTLDFGLARLLDIDETEXHADGSKPIKMMALSTLRPFTHQSD 904	
DB 837 DLNARNVLYSPNHYKTLDFGLARLLDIDETEXHADGSKPIKMMALSTLRPFTHQSD 896	
QY 905 VMSYGVVWELMFGAKPYDGIIPARBIPLDLKEGERLPQPCITIDVYIMVYKMDSDS 964	

FT DISULFID 219 221 BY SIMILARITY.
 FT DISULFID 232 240 BY SIMILARITY.
 FT DISULFID 248 260 BY SIMILARITY.
 FT DISULFID 251 260 BY SIMILARITY.
 FT DISULFID 264 291 BY SIMILARITY.
 FT DISULFID 285 307 BY SIMILARITY.
 FT DISULFID 311 326 BY SIMILARITY.
 FT DISULFID 329 333 BY SIMILARITY.
 FT DISULFID 356 515 BY SIMILARITY.
 FT DISULFID 510 523 BY SIMILARITY.
 FT DISULFID 526 535 BY SIMILARITY.
 FT DISULFID 539 555 BY SIMILARITY.
 FT DISULFID 558 571 BY SIMILARITY.
 FT DISULFID 562 579 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT DISULFID 595 617 BY SIMILARITY.
 FT DISULFID 620 628 BY SIMILARITY.
 FT DISULFID 624 636 BY SIMILARITY.
 FT MOD_RES 660 660 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
 FT MOD_RES 1092 1092 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1110 1110 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1172 1172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1197 1197 PHOSPHORYLATION (AUTO-) (MAJOR) (BY SIMILARITY).
 FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 19 19 C -> S (IN REF. 2).
 FT CONFLICT 539 539 C -> W (IN REF. 5).
 FT CONFLICT 991 991 L -> F (IN REF. 4).
 FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).
 SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46D2D2F5 CRC64;
 Query Match 46.2%; Score 3146; DB 1; Length 1210;
 Best Local Similarity 49.7%; Pred. No. 6,4e-159; Matches 635; Conservative 169; Mismatches 353; Indels 120; Gaps 24;

DB 419 AENLEIRGRKQGFSLAVANGANTISLGRSLKEISDGVITISGNNNLCYANTINK 478
 DB 484 QLEFRPHOALLHTARPDCEVSGEGLACGHCARGCMGPCTQCVCNCSQFLRGCEVSE 543
 DB 479 KLFGFNQKTKIMNRARCKCAVHVHVCNPLCSSSGCGPFRDVCSCNVSRGRCEVSE 538
 DB 544 CRVIGLREYVNAHCLPCHPECPONGSVTCFGEADQCVAAHYKDPFCVACPSG 603
 DB 539 CNLEGEEREFENECICHPBECLPQAMNITCTGRGPONCICQAHYIDGPHCVKCPAG 598
 DB 604 VKPDLSTVPIKPEPBEBAQCPCTNCHSCVDLDDKGPACQARSPLTSISAVGILL 663
 DB 599 IGENNNLT-VWKYADANVCHLCHANCCTYGCAGPGLQCEVWPSGPKLPSIATGVGILL 657
 DB 664 VYLVGVNFGI-LIKRQOKIRKYTRRRLLQETELVEPLTPSGAMPNOAMRLKETELRK 722
 DB 658 FIVV-VALGIGLPMRRRIIVKRTIRLLQRELEVEPLTPSGEARNQALRLKETEFK 716
 DB 723 VKVLGGAFGVYKGIWIPDGENVKI PVAIKVLRNTSPKANKELIDEAYVNAVGSPLY 782
 DB 717 IKVLGGAFGVYKGIWIPBEKVKI PVAIKELREATSPKANKELIDEAYVNAVSVNPHV 776
 DB 783 SHLGICITSTVQVLTQMLPVCCLLDHRENRGRGSGODLLWCMQIAGMSYLEDVRLV 842
 DB 777 CRLLGICITSTVQVLTQMLPVCCLLDVREHKNIGSOYLWCVQIAGMNYLEDRLV 836
 DB 843 HRLDARVNVKSPHVKITPGLARLLDIDETEVHADGKVPKIMMALESILRRFTHQ 902
 DB 837 HRLDARVNVKTPPHVKITPGLAKLGAEBKEHAEBGKPIKMALESILHRYTHQ 896
 DB 903 SVMSYGVTVWELMTFGAPYDGIPIAREIPDLLEKGERLPPICTIDVYIMVNCWMD 962
 DB 897 SVMSYGVTVWELMTFGSKPYDGIPIASDISILEKGERLPPICTIDVYIMVNCWMD 956
 DB 963 SECRREFELVSEFRMARDPQRFVYIO-NEDLGASPLDSTFFYSLEDDMGDLYDAE 1021
 DB 957 ADSRKREFELVSEFRMARDPQRFVYIO-NEDLGASPLDSTFFYSLEDDMGDLYDAE 1016
 DB 1022 EYLVPOGFEPCPPAPGAGWVHHRSSSTSGGDDLTGLPESEEARSPPLASEGA 1081
 DB 1017 EYLVPOGFEPCPPAPGAGWVHHRSSSTSGGDDLTGLPESEEARSPPLASEGA 1042
 DB 1082 GSDVEDGDLGMAKAGLSLPTHDSPLORYSEDPVLPSET--DGVAAPLTCSPQPEY 1139
 DB 1043 TSN-----NSTVACINRNSCRVXEDAPLQRYSSDPTGATEIDNIDAF-----PVPEY 1092
 DB 1140 VNQPDVRRPQPSREGRPLPARPAGATL-----ERPKTISPGKGVKQVPAFGAVENP 1194
 DB 1093 VNQ-----SVF-KRPAGSVONPVYHNPPLHPAPORDLHYQN--PHSNVAVGNP 1136
 DB 1195 EYLVNTAQ-----PTCLSSGFNSPALMIOKSHQSLNDPDIQDDFPFKETKNGIF 1237
 DB 1137 EYLVNTAQ-----PTCLSSGFNSPALMIOKSHQSLNDPDIQDDFPFKETKNGIF 1187
 DB 1238 KGTPTAENPEYLGIDVP 1254
 DB 1188 KG-PTAENAEYLRVAP 1203
 RESULT 6
 ERB4 HUMAN
 ID ERB4 HUMAN
 AC 015303;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)
 DE (p180erbB4) (Tyrosine kinase-type cell surface receptor HERR4).
 GN ERB4 OR HERR4
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

CC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM JM-A).
RC TISSUE=Breast carcinoma;
RX MEDLINE=93189574; PubMed=8383326;
RA Plowman G.D., Coloucou J.-M., Whitney G.S., Green J.M., Carlton G.W.,
RA Roy L., Neubauer M.G., Shoyab M.;
RT "A novel junctional membrane domain isoform of HER4/Erbb4. Isoform-specific
RT tissue distribution and differential processing in response to
RT phorbol ester.";
RL J. Biol. Chem. 272:26761-26768(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).
RC TISSUE=Petal brain;
RX MEDLINE=97476287; PubMed=9334463;
RA Elenius K., Cortas G., Paul S., Choi C.J., Rio C., Plowman G.D.,
RA Klagsbrun M.;
RT "A novel junctional membrane domain isoform of HER4/Erbb4. Isoform-specific
RT tissue distribution and differential processing in response to
RT phorbol ester.";
RL J. Biol. Chem. 272:26761-26768(1997).
RN [3]
RP INTERACTION WITH SNTB2.
RX MEDLINE=20202681; PubMed=10725395;
RA Garcia R.A., Vasudevan K., Buonanno A.;
RT "The neuroligin receptor ErbB-4 interacts with PDZ-containing proteins
RT at neuronal synapses.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:3596-3601(2000).
CC -1- FUNCTION: Specifically binds and is activated by neuroligin, NRG-
CC -2. NRG-3, heparin-binding EGF-like growth factor, betacellulin and
CC TNF α . Interaction with these factors induces cell differentiation.
CC Not activated by EGF, TGF- α , and amphiregulin.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: Homodimer or heterodimer with each of the other ERBB
CC receptors (potential). Interacts with the PDZ domain of the
CC synaptrophin SNTB2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Even-Alternative splicing; Named isoforms=2;
CC Comment=The 2 isoforms differ functionally in their response to
CC phorbol ester. Isoform JM-A is processed but not isoform JM-B.
CC So, they respectively represent cleavable and noncleavable
CC forms of the receptor. Both isoforms are expressed in the
CC cerebellum, but only the isoform JM-B is expressed in the
CC heart;
CC Name=JM-A;
CC IsoId=Q15303-1; Sequence=Displayed;
CC Name=JM-B;
CC IsoId=Q15303-2; Sequence=VSP_002895;
CC -1- TISSUE SPECIFICITY: Expressed at highest levels in brain, heart,
CC kidney, in addition to skeletal muscle, parathyroid, cerebellum,
CC placental, spleen, testis and breast. Lower levels in thymus,
CC lung, salivary gland, and pancreas.
CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine
CC residues.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/>
CC or send an email to license@isb.ch).
CC -----
DR EMBL, L07868, AB559446.1; -.
DR PIR, A47253, A47253.
DR HSPR, P11362, 1PGR.
DR GeneW, HGNC:3432, ERBB4.
DR MIM, 600543; -.
DR GO, GO:0005887, C:integral to plasma membrane; TAS.
DR GO, GO:0008283, P:cell proliferation; TAS.
DR GO, GO:0007275, P:development; TAS.

	KM	GO: GO:007048; P oncogenesis; TAS.
DR	InterPro:	IPIR006494; EGFR_L_domain.
DR	InterPro:	IPIR006211; Furin_like.
DR	InterPro:	IPIR006212; Furin_repeat.
DR	InterPro:	IPIR000719; Prot_kinase.
DR	InterPro:	IPIR001245; Tyr_kinase.
DR	InterPro:	IPIR004019; YLP motif.
DR	Pfam:	PF00757; Furin-like_1.
DR	Pfam:	PF00659; kinase_1.
DR	Pfam:	PF01050; Recep_L_domain; 2.
DR	PRINTS:	PRO0109; TYRKINASE.
DR	Prodom:	PDO00001; Prot_kinase; 1.
DR	SMART:	SMD0261; FU_5.
DR	SMART:	SMD0219; Tyrc_K_1.
DR	PROSITE;	PSO0107; PROTEIN KINASE_ATP; 1.
DR	PROSITE;	PSO0109; PROTEIN KINASE_TYR; 1.
DR	PROSITE;	PS50011; PROTEIN KINASE_DOM; 1.
KW	DataMembrane:	Glycoprotein_Multigene family; Receptor; Signal;
KW	Transferase:	Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KM	Alternative splicing.	
FT	SIGNAL	1 25 POTENTIAL.
FT	CHAIN	26 1308 RECEPTOR-PROTEIN-TYROSINE KINASE ERBB-4.
FT	DOMAIN	26 651 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	652 675 POTENTIAL.
FT	DOMAIN	676 1308 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	186 334 CYS-RICH.
FT	DOMAIN	496 633 CYS-RICH.
FT	DOMAIN	718 985 PROTEIN KINASE.
FT	NIP BIND	724 732 ACP (BY SIMILARITY).
FT	BINDING	751 751 ACP (BY SIMILARITY).
FT	ACT SITE	843 843 BY SIMILARITY.
FT	DISULFID	189 197 BY SIMILARITY.
FT	DISULFID	193 205 BY SIMILARITY.
FT	DISULFID	213 221 BY SIMILARITY.
FT	DISULFID	217 229 BY SIMILARITY.
FT	DISULFID	230 238 BY SIMILARITY.
FT	DISULFID	234 246 BY SIMILARITY.
FT	DISULFID	249 258 BY SIMILARITY.
FT	DISULFID	262 289 BY SIMILARITY.
FT	DISULFID	293 304 BY SIMILARITY.
FT	DISULFID	308 323 BY SIMILARITY.
FT	DISULFID	326 330 BY SIMILARITY.
FT	DISULFID	503 512 BY SIMILARITY.
FT	DISULFID	507 520 BY SIMILARITY.
FT	DISULFID	523 532 BY SIMILARITY.
FT	DISULFID	536 552 BY SIMILARITY.
FT	DISULFID	555 569 BY SIMILARITY.
FT	DISULFID	559 577 BY SIMILARITY.
FT	DISULFID	580 589 BY SIMILARITY.
FT	DISULFID	593 614 BY SIMILARITY.
FT	DISULFID	617 625 BY SIMILARITY.
FT	DISULFID	621 633 BY SIMILARITY.
FT	MOD_RES	1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1284 1284 MOD RES (BY SIMILARITY).
FT	CABOHAYD	138 138 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	174 174 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	181 181 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	253 253 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	358 358 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	410 410 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	473 473 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	495 495 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	548 548 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	576 576 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CABOHAYD	620 620 N-LINKED (GLCNAC...) (POTENTIAL).
FT	VARSPLIC	626 648 NGTSHDCIYYPMTGSTLPOHA -> IGSISIEDCIGLMK (In isoform JM-B).
SE	SEQUENCE	1308 AA; 146807 MM; 5EAAEB09E5DB8761 CRC64; /FTid=VSP 002895.

Query Match 44.1%; Score 3004.5; DB 1; Length 1308;
 Beet Local Similarity 45.6%; Pred. No. 2.1e-151;
 Matched 614; Conservative 183; Mismatches 375; Indels 175; Gaps 27;

9 WGLLALALPBGAA-----STVCGCTDMKRLRPASPTTHLDMLRHLVQGGVQVQGNLELY 64
 8 WYVSLVLAAGTVQPSDSQSCAGTENKLSLSLDEQYRALAKTYENCEVWVGNLEITS 67
 65 LPTNASLSPLQDIOEVQGYVLIAHQVROVPLQRLIRVGTQLFEDNVALAVLNDGPLN 124
 68 IHNNDLSPLRSVRETVGVVALNQFRLPLENLIIIGTKLYERVALAIFLNRKXG 127
 125 NTPPYTASPGGLRELQRLSLTEILKGYLIQNPOLCTQDTITLKQIDFHKNQALTLI 184
 128 NF-----GLOEGLKRLTEILNGVYVDQNKELCYADTITHMODIVRNWPMSNLTLV 178
 165 DNRBRACPCSPMCKSGRSMGSSSEDQSIFRTVAGG-ARCKPRLTDCCHBCAAG 243
 179 STNSSGCRCHKCTG-RCKGFTENHCOTLRITVCAEQDGRCTGYVSDCHRCAGG 237
 244 CTGPKHSDCLACLFPHNSGICELHCPALVTYNTDFESMNPREGRYTFGASCVTACPYN 303
 238 CSGPDTDCPACMNFDSGACVTCQPTVYVNPPTQLHNPAKATTYGAFCYKCKPHNF 297
 304 LSTDVGSCTLVCPRLHNOEYTAEDGTORCEKSKPCARVCYGLAMEHLREVAVTSANOE 363
 238 V-VDSGSCVRACPSRMEV-EENGIMCKPCPTDICPKACDGICTGSLMGAQVDSNDIK 355
 364 FAGCKKIPGSLAFLEPSFDGDPASNPAPQBPLOVFELEETGLVLSAMPDLSPLDS 423
 356 FNCTKINNLIFLVGTIHGDPNALIEALIDPELANTFRVREITGLANTQSPMPPNTDS 415
 424 VFQNLQVIRGRILHNGAVSLTLQSLGISWLGRLSLNELSGLALIHNTLCFVHTVMD 483
 416 VPSNLVTIGRVLVSGSLILKQGITSLQFQSLKLSAGNIYITDNLNLCYHTINNT 475
 484 QLFRRPHQALLHTANRPEDECVGEGIACQOLCARHGMGPRPQCYNCSQFLGQCEVB 543
 476 TLFSTINQRIVIRDNKKAENCTAEGMVCMHLCSSDCCMGPGPQCLSCRRFSRGRICBS 535
 544 CNVLGLPREVYNAHRCLECHPCOP-ONGSVTCRCPREADQCAVAAHYDPRPCVACRS 602
 536 CNLYDEEFEPENGSTCEBCDPCQCEMEGSLTCHBPGRPNCTKCSHFQDGPVCKCD 595
 603 GVKPDLSTYMPKWPDEBACQCPINTCTHSQVLDLDDKC-----PAQBRASPL 651
 596 GIGANSF--IKRYADPRDECHCHENCTQCGNCPISHDCTIYPTWGHTLPHAR---- 649
 652 TSIIISAVV--GILLVVLGVFGLIKRRQOKIRKYMRLLOETELVEPLTPSGAMPN 709
 650 TELIAGVIGGLFIVLTVGVAVVYRRRSIK-KQALRRFL-ETELVEPLTPSGAPNQ 707
 710 AQMRILKETELRKVKTLGSGAGCTYKGIWIPDGENVKIPVAIKVLRNTSPANKELID 769
 708 AQLRIKETELRKVKTLGSGAGCTYKGIWIPDGENVKIPVAIKVLRNTSPANKELID 767
 770 KAVYMAVGSPPYVSRLLGICLTSTVOLYTMPEYGLLDHVRNRRGLSODLLNMCMI 829
 768 EALINASMHPHLVRLGVCILSTIDVYOLMHPGCLLEVHKKONISQILLNMCVQI 827
 830 AKGMSYLEDVRLVHRDLAARNVLYKSPNHYKIDFGIARLLIDETEYHADGGKVIDIKM 889
 828 AKGMVLEERRRLVHRDLAARNVLYKSPNHYKIDFGIARLLIDETEYHADGGKVIDIKM 887
 890 ALESILRRFTQSDWVSGYTWELMTGAKPYDGIIPAREIDLLKGRRLPQPICTI 949
 888 ALBCIHYRKTQSDWVSGYTWELMTGAKPYDGIIPAREIDLLKGRRLPQPICTI 947
 950 DVMIMVKKCMIDSECRPRRELVSFSPNARBPQRFVYIONED-LGAPSLDSTYRSI 1008
 948 DVMIMVKKCMIDADSRPKFKELAAEBSNARBPQRYLYIQSDNRKCLSPENDSKFQNI 1007
 1009 LLEDMDGLVDABEYLVPQGFPCPDPAAGAMVHHRHSSTRSGGDLTLGLEPSEB 1068

DB 1008 LLEDMDGLVDABEYLVP-QATNIPP-----ITSRRLIDSNRS-----EIGSPPPA 1055
 QY 1069 EAPRS-----PLAP-SEGAGSDVFDGLGMAAGLQ 1099
 DB 1056 YTPMNGNQVTRDGGPAAEQSVVPYRAPTSNIPAPVAGQATAEFDSCCGTILRPV 1115
 QY 1100 SLPTHDPSPLOQYSDPFPVPLPS-----ETDGYAPRLTCSPPQPEYVQDPVPPSP 1152
 DB 1116 AHHVEDSSQYRSADPFFVAFERSPRGELDEGVTWTPMRDKPKQEYLVPE----- 1167
 QY 1153 REGPILPAPRAGATLERPKTSLSPKNGVQVYAFRGAIVENPEYLTPOGGAAPQHPPPA 1212
 DB 1168 -ENPFVSR-----KNGDLQ-----ALDNEEYHNASNG-----PPKA 1198
 QY 1213 -----FSPAFDNLVYMDOPPERGA--PPST 1236
 DB 1199 EDEYNEPELYNTPANTLCKAEYLNKLNILSMPEKAKADNDPDYVHNSLPPRSTLQHPDY 1258
 QY 1237 FKGPPT-----AENPEYL 1249
 DB 1259 LQESTKYFYKONGRIRPIVAENPEYL 1285

RESULT 7
 ERB4 RAT
 ID ERB4 RAT STANDARD; PRT; 1308 AA.
 AC 062956; 0942N7;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112).
 GN ERB4 OR TYRO-2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCB1_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=98221155; PubMed=9553078;
 RA Zhao Y.-Y., Sawyer D.R., Baliga R.R., Opel D.J., Han X.,
 Marchionni M.A., Kelly R.A.;
 RT "Neuregulin promote survival and growth of cardiac myocytes;
 RT Persistence of ErbB2 and ErbB4 expression in neonatal and adult
 RT ventricular myocytes.";
 RT J. Biol. Chem. 273:10261-10269(1998).
 RN (2)
 RP SEQUENCE OF 848-901 FROM N.A.
 RC TISSUE=Sciatic nerve;
 RX MEDLINE=9122560; PubMed=2025425;
 RA Lai C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially
 RT expressed in the vertebrate nervous system.";
 RT Neuron 6:691-704(1991).
 RN (3)
 RP SEQUENCE OF 1031-1198 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Spinal cord;
 RX MEDLINE=97484212; PubMed=9030624;
 RA Carroll S.L., Miller M.L., Frohner P.W., Kim S.S., Corbett J.A.;
 RT "Expression of neuregulins and their putative receptors, ErbB2 and
 RT ErbB3, is induced during Wallerian degeneration.";
 RT J. Neurosci. 17:1642-1659(1997).
 CC -1- FUNCTION: Specifically binds and is activated by neuregulins, NRG-
 CC 2, NRG-3, heparin-binding EGF-like growth factor, betacellulin and
 CC NTAK. Interaction with these factors induces cell differentiation.
 CC Not activated by EGF, TGF-A, and amphiregulin (by similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: Homodimer or heterodimer with each of the other ERBB
 CC receptors. Interacts with the PDZ domain of the synrophin SHB2
 CC (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: Preferentially expressed in the developing
 CC nervous system. Exhibits distinct and highly regionalized patterns
 CC of expression in the adult brain, where it is mainly found in the
 CC reticular nucleus of the thalamus. Very low levels in kidney, and
 CC heart.
 CC -1- PTH: Ligand-binding increases phosphorylation on tyrosine
 CC residues (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 CC EMBL, AF041838, AAD08899.1, -.
 CC EMBL, U52531, AAC53051.1, -.
 CC PIR, P70184, P70184.
 CC HSSP, P11362, 1FGK.
 CC InterPro, IPR000494, EGFR_L_domain.
 CC InterPro, IPR006211, Furin-like.
 CC InterPro, IPR006212, Furin-repeat.
 CC InterPro, IPR007119, Prot_kinase.
 CC InterPro, IPR001245, Tyr_kinase.
 CC InterPro, IPR004019, YLP_motif.
 CC Pfam, PF00757, Furin-like, 1.
 CC Pfam, PF00069, pkinase, 1.
 CC Pfam, PF01030, Recep_L_domain, 2.
 CC Pfam, PF02757, YLP, 2.
 CC Pfam, PF00001, Prot_kinase, 1.
 CC SMART, SMO0261, FU, 5.
 CC SMART, SMO0219, Tyrc, 1.
 CC PROSITE, PS00107, PROTEIN_KINASE_ATP, 1.
 CC PROSITE, PS0011, PROTEIN_KINASE_DOM, 1.
 CC PROSITE, PS00109, PROTEIN_KINASE_TYR, 1.
 CC Trnsmembrane, Glycoprotein; Multigene family; Receptor; Signal;
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
 CC SIGNAL 1 25
 FT CHAIN 26 1308 RECEPTOR PROTEIN-TYROSINE KINASE EMBB-4.
 FT DOMAIN 26 651 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 652 675 POTENTIAL.
 FT DOMAIN 186 1308 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 496 633 CYS-RICH.
 FT DOMAIN 718 985 CYS-RICH.
 FT NP_BIND 724 732 ATP (BY SIMILARITY).
 FT BINDING 751 751 ATP (BY SIMILARITY).
 FT ACT_SITE 843 843 BY SIMILARITY.
 FT DISULFID 189 197 BY SIMILARITY.
 FT DISULFID 193 205 BY SIMILARITY.
 FT DISULFID 213 221 BY SIMILARITY.
 FT DISULFID 217 229 BY SIMILARITY.
 FT DISULFID 230 238 BY SIMILARITY.
 FT DISULFID 244 258 BY SIMILARITY.
 FT DISULFID 252 289 BY SIMILARITY.
 FT DISULFID 293 304 BY SIMILARITY.
 FT DISULFID 308 323 BY SIMILARITY.
 FT DISULFID 326 330 BY SIMILARITY.
 FT DISULFID 503 512 BY SIMILARITY.
 FT DISULFID 520 532 BY SIMILARITY.
 FT DISULFID 533 552 BY SIMILARITY.
 FT DISULFID 536 552 BY SIMILARITY.
 FT DISULFID 555 569 BY SIMILARITY.
 FT DISULFID 559 577 BY SIMILARITY.
 FT DISULFID 580 589 BY SIMILARITY.
 FT DISULFID 593 614 BY SIMILARITY.
 FT DISULFID 617 625 BY SIMILARITY.
 FT DISULFID 621 633 BY SIMILARITY.
 FT MOD_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1168 1168

FT MOD_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 358 358 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 473 473 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 495 495 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 548 548 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 576 576 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 1062 1062 S -> N (IN REF. 3).
 FT CONFLICT 1080 1082 PYT -> SYR (IN REF. 3).
 SQ SEQUENCE 1308 AA; 146957 MW; D944B096A08B41 CRC64;
 Query Match 43.9%; Score 2889; DB 1; Length 1308;
 Beet Local Similarity 46.2%; Pred. No. 1.4e-150;
 Matches 612; Conservative 194; Mismatches 402; Indels 118; Gaps 27;
 1 MEA-ALCRWGLL--ALLPGAASTQVCTGDMRLRASPETHLDMLRHLYOGCQVQ 57
 1 MKATGLWVMSGLVAARTVQPSASQVCACTENLSSLSDEQGYRALRYKYECEVM 60
 58 GLEITVLPNASLSEFLDIOEVQVLIANQVQVPLQRLIRVGTQLFEDNVALAVL 117
 61 GLEITTSIEHNRDLSEFLSIREVGYVALNQFVPLPLENRIIRGTLYEDRYALAI 120
 118 DNGDPLNNTTPTVQASPGGLARELOASLTETLKGVULQORNQOLYODPILTKDIFHKRN 177
 121 LNYRKGNF-----GLDELGLKNTLELNGVYVDQNFELCYADITHQDVRNPW 171
 178 QALATLIDITNRSRACHPCSPCKSGRCSGESEDCSLTRVYACAGGC-ARCKGPLPTDCC 236
 172 PSNMTLVSTIGSSGGRCHKSTG-RCKGPTENHQTILTRYCAQDCRCKGPIVSDCC 230
 237 HQCAAGTGPGRHSDCLALPHNHSIGELHCPALVTNTDTPESMPNDEGRYTFGASCV 296
 231 HRECAAGSGPRDTPCFACNMNFDSGACVTCQPPQFVNPPTFQLEHNFNATYGAFCV 290
 237 TACPNVISTVGSCTLVCPILNQVTAEDQREBKSKPCARCYGGMELHREVRVAV 356
 291 KCPNPNFV-VSSSCVRACPSKKEV-BENGKMKKPCPTDICKPKADGIGTSLMSAQTV 348
 357 TSANIOFAGCKKIFGSLAFIPESPDGPASNTAIPLOPELOVFEETLEITGYLSAMP 416
 349 DSNIDKINCKINGNLIFLVYGHGDPYNAIDAIDEKLANVFRVRIITGFLNQTMP 408
 417 DSLPDLVYFONLOVIRGRILHNGAVSLTLOGLISMLGSLRSLRELSGLALIHNTLCP 476
 409 PMTDPVSFVSNLVITIGRVLVSGLLILKQGGITSLQFOSIKETISAGNIYITDNNLCV 468
 477 VHTVWDOLFRRPHALHTANRPREDCEYBGLACHOLCARHCHNGPRTQCNCSQFLR 536
 469 YHTIMWTLFTSVNRIYIDNRRAENCTABEMVNCNHLCSNDGCMGPGDQCLSCRFER 528
 537 GGECEVEGRVLOGREVENARHCLPCHPECP-ONGSVTCGPEADOCVACAHYDPPF 595
 529 GKICIESCNLDGEREFENGISCYECDSQCEKMDGLTCHGPRPDNTCKSHPRDGN 568
 596 CVARCPGKPDLSYMPWKPEDEBGAOCPCINCTHSCVDLDDKGC-----PA 644
 589 CYEKCPDVLQGANSE--IFKYADQRECHPCPNCTOGCNGPSTSHCIYYPTGSHLQ 646
 645 EGRAPPLSIISAV--GILVVYGVVPGIIXKROQKIRKYTRMRLLQETELVPLRP 702
 647 HAR---TPLLAAGVIGGFLIVIALTFAYVVRKRSIK-KRALRFL-ETELVPLRP 700
 703 SGAMPNOMKRLKTELRKYVLGSSGAFGYKGIWIPDGNVKKIPVAILVRENTSPK 762
 701 SGTANQOAKRLKETELKRYVVLGSSGAFGYKGIWVEGELTKIPVAILKINETGPK 760
 763 ANKEILDEAYVAGVSPVSRLLGICLTSTVOLVTQMLPYGCLLDHVENRGRGLSQDL 822

```

Db 761 ANVEPMDBALIMASVDHPLVLLGLVCLSPITQVLTQIMPHOCLLEYEHENDONIGSOL 820
Cc 823 LNMCGQIKNGMSYLDVRLVHEDLAARNVLYKSPHNVKTTDGLARLIDIDETRYHADGG 882
Cc 821 LNMCGQIKNGMYLERRLVHEDLAARNVLYKSPHNVKTTDGLARLIDIDETRYHADGG 880
Cc 863 KVPIMKMLLESILRRRFTHQSDVMSYGVTVWELMTFGAKPYDQIPAREIPDLLEKGERLP 942
Cc 861 KMPIMKMLLECITHYKFTHQSDVMSYGVTVWELMTFGKPYDQIPAREIPDLLEKGERLP 940
Cc 943 OPTICTIVYMMVKCMWIDSECRPRFELVSEPMARDPQRFVYIQED-LGASPLD 1001
Cc 941 OPTICTIVYMMVKCMWIDSECRPRFELVSEPMARDPQRFVYIQED-LGASPLD 1000
Cc 1002 STFYRLLEDMDGLVDAEELVLPQGGFFCPDP-----ARG 1038
Cc 1001 SKFFQNLDEEDLEMDAEELVLP-QAFNIPPIYTSRTRIDSNRSEIGHSPPRATYPM 1059
Cc 1039 AGGVVHRRSSSTRSGGDLTLGLEPSEBEAPRSPLASEGSDVDPDGLMGAKGL 1098
Cc 1060 SGSGVYDGGGATQGG--MEMPTATATSTIPAPVA--QCATLMEFDDSCNGTLRKP 1114
Cc 1099 QSLPHDPSPLQRYSDPVPPLPS-----ETDGVAPLTCSPQEPYVNDVAPRPPS 1151
Cc 1115 VVPHVQSDSSTQRYSDPVPAPERNPRAELDEEGTMTMHDKPKQXYLNPAENPFVSR 1174
Cc 1152 PREGPLPARPAGATLERPK---TLSPGKGVVQVF-----AFGAVENPEYLTP 1199
Cc 1175 RINGLOA-----LDNPEYTSASGPK---ABEYVNEPLYLNTFNALGNMEYMKM 1224
Cc 1200 QGGAPOHPPEPAPFNLYWDODPERGA--PPSTFKGPT-----A 1243
Cc 1225 SLASVPE-----KAKKAFDNPDMYHNSLPRSTGLQPDVLTQGYSTYFYKNGRIPIYA 1279
Cc 1244 ENPEYL 1249
Cc 1280 ENPEYL 1285

```

RESULT 8

```

XMRK_X1PMA STANDARD; PRT: 1167 AA.
ID_XMRK_X1PMA
AC P1338;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Melanoma receptor protein-tyrosine kinase precursor (EC 2.7.1.112).
GN XMRK OR TV.
OS Xiphophorus maculatus (Southern platyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Percomorphia; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_TaxID=8083;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90015140; PubMed=2797166;
RA Witbrodt J., Adam D., Maltzschek B., Maueier W., Raulf F.,
RA Telling A., Robertson S.M., Schartl M.;
RT "Novel putative receptor tyrosine kinase encoded by the melanoma-
RT inducing Tu locus in Xiphophorus";
RL Nature 341:415-421 (1998).
RP [2]
RP REVISION TO 515.
RA Schartl M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLE RECEPTOR WITH PROTEIN-TYROSINE KINASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DISEASE: INVOLVED IN PIGMENT CELLS MALIGNANT MELANOMAS.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

```

```

Cc -----
Cc This SWISS-Prot entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (see http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL: X16891; CAA34770.2;
Cc PIR: S06142; S06142.
Cc HSP, P1362, 1REK.
Cc InterPro: IPR00494; EGFR_L_domain.
Cc InterPro: IPR006211; Furin-like.
Cc InterPro: IPR006212; Furin repeat.
Cc InterPro: IPR00719; Prot_kinase.
Cc InterPro: IPR01245; Tyr_kinase.
Cc Pfam: PF00757; Furin-like; 1.
Cc Pfam: PF0069; Kinase; 1.
Cc Pfam: PF01030; Recep_L_domain; 2.
Cc PRINTS: PS0109; TYRKINASE.
Cc ProDom: PD00001; Prot_kinase; 1.
Cc SMART: SM00261; Fu; 5.
Cc SMART: SM00219; Tyrc; 1.
Cc PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
Cc PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
Cc Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
Cc Tyrosine-protein kinase; ATP-binding; phosphorylation; Proto-oncogene.
Cc SIGNAL
Cc CHAIN
Cc 26 1167
Cc 26 642
Cc 643 665
Cc 666 1167
Cc 710 977
Cc 716 743
Cc 743 743
Cc 835 835
Cc 835 835
Cc 195 204
Cc 199 212
Cc 220 228
Cc 224 236
Cc 237 245
Cc 241 253
Cc 256 265
Cc 269 296
Cc 300 311
Cc 315 330
Cc 333 337
Cc 504 513
Cc 508 521
Cc 524 533
Cc 537 553
Cc 556 569
Cc 560 577
Cc 593 615
Cc 618 626
Cc 622 634
Cc 634 642
Cc 114 114
Cc 144 144
Cc 201 201
Cc 356 356
Cc 365 365
Cc 398 398
Cc 417 417
Cc 501 501
Cc 576 576
Cc 621 621
Cc 1167 AA; 129934 MW; 479380749DCID55A CRC64;

```

Query Match 40.0%; Score 2724.5; DB 1; Length 1167;
Best Local Similarity 45.5%; Pred. No. 1.1e-136;

Matches 577; Conservative 164; Mismatches 389; Indels 137; Gaps 27;

4 AALCRWGLLALLPFGAAST---QVCTGTDMLRLPAPETHLDMRLHLYOCQVQVGN 59
 8 AALLQ--LLLVLSIRCSCTDPDRKVCQOSTNQMTM---LDNHYLMKKMVGSCNVVLEN 62
 60 LELTYLPLTNASISPLDIOEVOGYVLIANOVROVPLQRLIVRGTFQEDVYALALDN 119
 63 LEITVQEQDLSPLQIEVGVYVLIANNEVSTIPLVNLRLIRGQNLVEGNFTLLVMSN 122
 120 GDP LNNTPTVTCASFGCLRELQRLSTELLKGVLIQRPOLCYODTILMKDIFHKNNL 179
 123 YOK-PPSSP--DVYGVGKQQLSNLTLLSGVKSHPNPLLCNVETIMMDI VDKTSP 179
 180 ALLTIDTRNSRACHPCSPWCKGSRGWSSESDCQSLTRTVCAAGC--ARCKGPLEPTDCHE 238
 180 TMLLPLHAFBRGQCKDHCNVGSCAPRPGHCQKTKLLCAEQCRNRKRGKPLDCCNB 239
 239 OCAAGCTGPKHSDCLACLPHNSGICELACPALVTYNTDTFESMPNREGRYTFGASCVA 298
 240 HCAGGCTGPRATDCLACRDFNDGCTKOTCPPKTYDIYSHQVVDNPNIKYTFGAACVKE 299
 299 CRYNTLSTDVGSCTVCPLEHNOETVABDTQRCCKSKCAVCYGLGMEHLREVAATVS 358
 300 CPSNTVYVTE--GACVNSCSAGMLEVD--ENKRSCKPCDGVCPKVCDDIGISLNTIANS 357
 359 ANIOBPAGCKKIFGSLAFPESSPDGPPASNTAPLOPELOVETLEITGYLYISAMPDS 418
 358 TNIRSFNSTKINGDILINRNSFBDPHYKIGTMDBEHMLTLYVEITGYLYIMMPEFN 417
 419 LPLDSVFOQLQVIRGRILHNGAYS--LTLOGLSIWLGLSLBELSGSLIHHNTHLCFV 477
 418 MTSLSVFOQLLEIRGRTTSRGSFVYGVVRHQLGLASLKEVSGANVILNKTLQLRYA 477
 478 KTVPPDQLFRMHQALHTAARPEDECVBGLAQHLCARGHCWGGEPFQCNCQSPFLRG 537
 478 NTINNRRLRSESDQSI EYDART-----ENOTCNNESEBDCWGMGPFMCVSCILAVDNG 530
 538 QECVEBCRYLQGLPREYVNAARCLPCHPECOPONGSVTFGEADQCVACAHYKDPFCV 597
 531 GRVNASCNLQGBPREAOVDGRVCQCHQCLVOTBSLTITCYGEPANCSSAHFQDPQCI 590
 598 ACPSPGVKEDLSYMPIWKPDEBEGACQPCPINCTHSCVDLDDKCAPAORASPLTSIISA 657
 591 PCPFGILGDPDTL-IMKYADKMGQCPHQNCTQCSGPGISGCGD--IVSHSLAVGL 648
 658 VVGLILVVLGVVFGILIKRQOKIRKYMRLQSTELVEPLTPSGAMPNQORILKE 717
 649 VSGLLITVIVALLIVLLRRRIK--RKRTIRCLQEKELVEPLTPSGQAPNOAFILIKE 707
 718 TELRKVKVLSGSAFGVYVYKGIPIPOGENYKIPALIVLRENTSPKANKELDEAYYMACV 777
 708 TEPKDQVLSGSAFGVYVYKGMNPODENRIRIPALIVLRENTSPKANKELDEAYYMACV 767
 778 GSPYVSRLLGICLTSTVQVLTQMPYGLLDHVENRGLSGQDLNMCQIAKMSYLE 837
 768 DHPHCRLRIGLICLTSAVOLVTOQMPYGLLDYVROHERICQGLNMCVQIAKMSYLE 827
 838 DVRLVHRDLAARVYVKSPPHYKITDPLGLAKLIDIDETVHADGKVPYIKMALESILR 897
 828 EHLVHRDLAARVYVKSPPHYKITDPLGLAKLIDIDETVHADGKVPYIKMALESILR 887
 898 RFTHOSDVSYGVTVLMTFGAKPYDGI-PAREIPLEKGERLPPPICTIDVYIMYK 957
 888 TYTHOSDVSYGVTVLMTFGSKPYDGI-PAKEIASVLENGERLPPPICTIEIVMIIIK 947
 958 CMVIDSECPREFRELVSFSRMAARDQRFVYVIONEDLGPASPLDSTFYRSLLEDMDGL 1017
 948 CMVIDSSPRPRELVSFSRMAARDQRFVYVIONEDLGPASPLDSTFYRSLLEDMDGL 1002
 1018 VNAEVLIVQGGFCFCDPAPAGAGCAVHHRRSSTSGGDLTLGLEPSEEAAPRPLAP 1077
 1003 VNAEVLIVQGGFCFCDPAPAGAGCAVHHRRSSTSGGDLTLGLEPSEEAAPRPLAP 1025

1078 SEGASDVFDGLGMAKGLQSLPETHDPSPLQRYSEDPV--PLPSETDGVVAPLTCSPO 1136
 1026 PLGH-----PVRENSITLRLNISPTQVLALEKDLDDG----- 1056
 1137 PEYVQPPVPRPOP-----PSPRE-----GPLP--AARPAQTLEPRKTLSPGKGVKD 1183
 1057 -EYVQPPSETSSRLSDIYNPYEDLTGCGVPSLSSQEAETNFSRPEYLTNTNONS----- 1112
 1184 VFAPGAVENPEYLTPOGGAAPQHPAPFAPFQNLVYMDODPPERGAPSTFKCTPTA 1243
 1113 PLVSSGSHDDPEY---QAG-----YQAAF-----LPQTGALTNGMFLPPA 1150
 1244 ENPEYLG 1250
 1151 ENLEYLG 1157

RESULT 9
 ERB3 HUMAN
 ID ERB3 HUMAN . STANDARD; PRT; 1342 AA.
 AC P21860;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)
 DE (c-erbB3) (Tyrosine kinase-type cell surface receptor HER3).
 GN ERB3 OR HER3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=90083234; PubMed=2687875;
 RA Kraus M.H., Issing W., Miki T., Popescu N.C., Aaronson S.A.;
 RT "Isolation and characterization of ERB3, a third member of the
 RT ERBB/epidermal growth factor receptor family: evidence for
 RT overexpression in a subset of human mammary tumors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:9193-9197(1989).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=90311312; PubMed=2164210;
 RA Plozman G.D., Whitney G.S., Neubauer M.G., Green J.M., McDonald V.L.,
 RA Todaro G.J., Shoyab M.;
 RT "Molecular cloning and expression of an additional epidermal growth
 RT factor receptor-related gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:4905-4909(1990).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=placenta;
 RX MEDLINE=93282822; PubMed=7685162;
 RA Katch M., Yazaki Y., Sugimura T., Terada M.;
 RT "c-erbB3 gene encodes secreted as well as transmembrane receptor
 RT tyrosine kinase.";
 RL Biochem. Biophys. Res. Commun. 192:1189-1197(1993).
 CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTAK.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
 CC (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG FORM) AND
 CC SECRETED (SHORT FORM).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=long form;
 CC IsoId=P21860-1; Sequence=Displayed;
 CC Name=2; Synonyms=short form;
 CC IsoId=P21860-2; Sequence=VSP 002893, VSP 002894;
 CC -1- TISSUE SPECIFICITY: EPITHELIAL TISSUES AND BRAIN.
 CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
 CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES

CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
 CC PHOSPHATIDYLINOSITOL 3-KINASE (BY SIMILARITY).
 CC -1- DISEASE: OVEREXPRESSED IN A SUBSET OF HUMAN MAMMARY TUMORS.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ch/announce/>
 CC or send an email to license@ebi.ebi.ac.uk).
 CC -----
 CC EMBL, M29366; AAA5790.1; -;
 CC EMBL, M34309; AAA5979.1; -;
 CC EMBL, S61953; AAB26935.1; -;
 CC PIR, A36223; A36223.
 CC PIR, JH0803; JH0803.
 CC PDB, 1M6B; 23-AUG-02.
 CC GeneW, HGNC:3431; ERBB3.
 CC MIM, 190151; -;
 CC GO, GO:0005887; C: integral to plasma membrane; TAS.
 CC GO, GO:0005906; E: epidermal growth factor receptor activity; TAS.
 CC GO, GO:0006468; P: protein amino acid phosphorylation; TAS.
 CC InterPro, IPR000494; EGFR_L domain.
 CC InterPro, IPR006211; Furin-like.
 CC InterPro, IPR006212; Furin repeat.
 CC InterPro, IPR000719; Prot Kinase.
 CC InterPro, IPR001245; Tyr_Kinase.
 CC Pfam, PF00757; Furin-like; 1.
 CC Pfam, PF00069; pkinase; 1.
 CC Pfam, PF01030; Recep_L domain; 2.
 CC PRINTS, PR00109; TYRKINASE.
 CC ProDom, PD000001; Proc_kinase; 1.
 CC SMART, SM00261; PU.5.
 CC SMART, SM00219; TyKc; 1.
 CC PROSITE, PS00107; PROTEIN KINASE ATP, FALSE NEG.
 CC PROSITE, PS00109; PROTEIN KINASE TYR, FALSE NEG.
 CC PROSITE, PS50011; PROTEIN KINASE DOM; 1.
 CC Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 CC Alternative splicing; 3D-structure.
 CC K1
 CC K2
 CC K3
 CC K4
 CC K5
 CC K6
 CC K7
 CC K8
 CC K9
 CC K10
 CC K11
 CC K12
 CC K13
 CC K14
 CC K15
 CC K16
 CC K17
 CC K18
 CC K19
 CC K20
 CC K21
 CC K22
 CC K23
 CC K24
 CC K25
 CC K26
 CC K27
 CC K28
 CC K29
 CC K30
 CC K31
 CC K32
 CC K33
 CC K34
 CC K35
 CC K36
 CC K37
 CC K38
 CC K39
 CC K40
 CC K41
 CC K42
 CC K43
 CC K44
 CC K45
 CC K46
 CC K47
 CC K48
 CC K49
 CC K50
 CC K51
 CC K52
 CC K53
 CC K54
 CC K55
 CC K56
 CC K57
 CC K58
 CC K59
 CC K60
 CC K61
 CC K62
 CC K63
 CC K64
 CC K65
 CC K66
 CC K67
 CC K68
 CC K69
 CC K70
 CC K71
 CC K72
 CC K73
 CC K74
 CC K75
 CC K76
 CC K77
 CC K78
 CC K79
 CC K80
 CC K81
 CC K82
 CC K83
 CC K84
 CC K85
 CC K86
 CC K87
 CC K88
 CC K89
 CC K90
 CC K91
 CC K92
 CC K93
 CC K94
 CC K95
 CC K96
 CC K97
 CC K98
 CC K99
 CC K100
 CC K101
 CC K102
 CC K103
 CC K104
 CC K105
 CC K106
 CC K107
 CC K108
 CC K109
 CC K110
 CC K111
 CC K112
 CC K113
 CC K114
 CC K115
 CC K116
 CC K117
 CC K118
 CC K119
 CC K120
 CC K121
 CC K122
 CC K123
 CC K124
 CC K125
 CC K126
 CC K127
 CC K128
 CC K129
 CC K130
 CC K131
 CC K132
 CC K133
 CC K134
 CC K135
 CC K136
 CC K137
 CC K138
 CC K139
 CC K140
 CC K141
 CC K142
 CC K143
 CC K144
 CC K145
 CC K146
 CC K147
 CC K148
 CC K149
 CC K150
 CC K151
 CC K152
 CC K153
 CC K154
 CC K155
 CC K156
 CC K157
 CC K158
 CC K159
 CC K160
 CC K161
 CC K162
 CC K163
 CC K164
 CC K165
 CC K166
 CC K167
 CC K168
 CC K169
 CC K170
 CC K171
 CC K172
 CC K173
 CC K174
 CC K175
 CC K176
 CC K177
 CC K178
 CC K179
 CC K180
 CC K181
 CC K182
 CC K183
 CC K184
 CC K185
 CC K186
 CC K187
 CC K188
 CC K189
 CC K190
 CC K191
 CC K192
 CC K193
 CC K194
 CC K195
 CC K196
 CC K197
 CC K198
 CC K199
 CC K200
 CC K201
 CC K202
 CC K203
 CC K204
 CC K205
 CC K206
 CC K207
 CC K208
 CC K209
 CC K210
 CC K211
 CC K212
 CC K213
 CC K214
 CC K215
 CC K216
 CC K217
 CC K218
 CC K219
 CC K220
 CC K221
 CC K222
 CC K223
 CC K224
 CC K225
 CC K226
 CC K227
 CC K228
 CC K229
 CC K230
 CC K231
 CC K232
 CC K233
 CC K234
 CC K235
 CC K236
 CC K237
 CC K238
 CC K239
 CC K240
 CC K241
 CC K242
 CC K243
 CC K244
 CC K245
 CC K246
 CC K247
 CC K248
 CC K249
 CC K250
 CC K251
 CC K252
 CC K253
 CC K254
 CC K255
 CC K256
 CC K257
 CC K258
 CC K259
 CC K260
 CC K261
 CC K262
 CC K263
 CC K264
 CC K265
 CC K266
 CC K267
 CC K268
 CC K269
 CC K270
 CC K271
 CC K272
 CC K273
 CC K274
 CC K275
 CC K276
 CC K277
 CC K278
 CC K279
 CC K280
 CC K281
 CC K282
 CC K283
 CC K284
 CC K285
 CC K286
 CC K287
 CC K288
 CC K289
 CC K290
 CC K291
 CC K292
 CC K293
 CC K294
 CC K295
 CC K296
 CC K297
 CC K298
 CC K299
 CC K300
 CC K301
 CC K302
 CC K303
 CC K304
 CC K305
 CC K306
 CC K307
 CC K308
 CC K309
 CC K310
 CC K311
 CC K312
 CC K313
 CC K314
 CC K315
 CC K316
 CC K317
 CC K318
 CC K319
 CC K320
 CC K321
 CC K322
 CC K323
 CC K324
 CC K325
 CC K326
 CC K327
 CC K328
 CC K329
 CC K330
 CC K331
 CC K332
 CC K333
 CC K334
 CC K335
 CC K336
 CC K337
 CC K338
 CC K339
 CC K340
 CC K341
 CC K342
 CC K343
 CC K344
 CC K345
 CC K346
 CC K347
 CC K348
 CC K349
 CC K350
 CC K351
 CC K352
 CC K353
 CC K354
 CC K355
 CC K356
 CC K357
 CC K358
 CC K359
 CC K360
 CC K361
 CC K362
 CC K363
 CC K364
 CC K365
 CC K366
 CC K367
 CC K368
 CC K369
 CC K370
 CC K371
 CC K372
 CC K373
 CC K374
 CC K375
 CC K376
 CC K377
 CC K378
 CC K379
 CC K380
 CC K381
 CC K382
 CC K383
 CC K384
 CC K385
 CC K386
 CC K387
 CC K388
 CC K389
 CC K390
 CC K391
 CC K392
 CC K393
 CC K394
 CC K395
 CC K396
 CC K397
 CC K398
 CC K399
 CC K400
 CC K401
 CC K402
 CC K403
 CC K404
 CC K405
 CC K406
 CC K407
 CC K408
 CC K409
 CC K410
 CC K411
 CC K412
 CC K413
 CC K414
 CC K415
 CC K416
 CC K417
 CC K418
 CC K419
 CC K420
 CC K421
 CC K422
 CC K423
 CC K424
 CC K425
 CC K426
 CC K427
 CC K428
 CC K429
 CC K430
 CC K431
 CC K432
 CC K433
 CC K434
 CC K435
 CC K436
 CC K437
 CC K438
 CC K439
 CC K440
 CC K441
 CC K442
 CC K443
 CC K444
 CC K445
 CC K446
 CC K447
 CC K448
 CC K449
 CC K450
 CC K451
 CC K452
 CC K453
 CC K454
 CC K455
 CC K456
 CC K457
 CC K458
 CC K459
 CC K460
 CC K461
 CC K462
 CC K463
 CC K464
 CC K465
 CC K466
 CC K467
 CC K468
 CC K469
 CC K470
 CC K471
 CC K472
 CC K473
 CC K474
 CC K475
 CC K476
 CC K477
 CC K478
 CC K479
 CC K480
 CC K481
 CC K482
 CC K483
 CC K484
 CC K485
 CC K486
 CC K487
 CC K488
 CC K489
 CC K490
 CC K491
 CC K492
 CC K493
 CC K494
 CC K495
 CC K496
 CC K497
 CC K498
 CC K499
 CC K500
 CC K501
 CC K502
 CC K503
 CC K504
 CC K505
 CC K506
 CC K507
 CC K508
 CC K509
 CC K510
 CC K511
 CC K512
 CC K513
 CC K514
 CC K515
 CC K516
 CC K517
 CC K518
 CC K519
 CC K520
 CC K521
 CC K522
 CC K523
 CC K524
 CC K525
 CC K526
 CC K527
 CC K528
 CC K529
 CC K530
 CC K531
 CC K532
 CC K533
 CC K534
 CC K535
 CC K536
 CC K537
 CC K538
 CC K539
 CC K540
 CC K541
 CC K542
 CC K543
 CC K544
 CC K545
 CC K546
 CC K547
 CC K548
 CC K549
 CC K550
 CC K551
 CC K552
 CC K553
 CC K554
 CC K555
 CC K556
 CC K557
 CC K558
 CC K559
 CC K560
 CC K561
 CC K562
 CC K563
 CC K564
 CC K565
 CC K566
 CC K567
 CC K568
 CC K569
 CC K570
 CC K571
 CC K572
 CC K573
 CC K574
 CC K575
 CC K576
 CC K577
 CC K578
 CC K579
 CC K580
 CC K581
 CC K582
 CC K583
 CC K584
 CC K585
 CC K586
 CC K587
 CC K588
 CC K589
 CC K590
 CC K591
 CC K592
 CC K593
 CC K594
 CC K595
 CC K596
 CC K597
 CC K598
 CC K599
 CC K600
 CC K601
 CC K602
 CC K603
 CC K604
 CC K605
 CC K606
 CC K607
 CC K608
 CC K609
 CC K610
 CC K611
 CC K612
 CC K613
 CC K614
 CC K615
 CC K616
 CC K617
 CC K618
 CC K619
 CC K620
 CC K621
 CC K622
 CC K623
 CC K624
 CC K625
 CC K626
 CC K627
 CC K628
 CC K629
 CC K630
 CC K631
 CC K632
 CC K633
 CC K634
 CC K635
 CC K636
 CC K637
 CC K638
 CC K639
 CC K640
 CC K641
 CC K642
 CC K643
 CC K644
 CC K645
 CC K646
 CC K647
 CC K648
 CC K649
 CC K650
 CC K651
 CC K652
 CC K653
 CC K654
 CC K655
 CC K656
 CC K657
 CC K658
 CC K659
 CC K660
 CC K661
 CC K662
 CC K663
 CC K664
 CC K665
 CC K666
 CC K667
 CC K668
 CC K669
 CC K670
 CC K671
 CC K672
 CC K673
 CC K674
 CC K675
 CC K676
 CC K677
 CC K678
 CC K679
 CC K680
 CC K681
 CC K682
 CC K683
 CC K684
 CC K685
 CC K686
 CC K687
 CC K688
 CC K689
 CC K690
 CC K691
 CC K692
 CC K693
 CC K694
 CC K695
 CC K696
 CC K697
 CC K698
 CC K699
 CC K700
 CC K701
 CC K702
 CC K703
 CC K704
 CC K705
 CC K706
 CC K707
 CC K708
 CC K709
 CC K710
 CC K711
 CC K712
 CC K713
 CC K714
 CC K715
 CC K716
 CC K717
 CC K718
 CC K719
 CC K720
 CC K721
 CC K722
 CC K723
 CC K724
 CC K725
 CC K726
 CC K727
 CC K728
 CC K729
 CC K730
 CC K731
 CC K732
 CC K733
 CC K734
 CC K735
 CC K736
 CC K737
 CC K738
 CC K739
 CC K740
 CC K741
 CC K742
 CC K743
 CC K744
 CC K745
 CC K746
 CC K747
 CC K748
 CC K749
 CC K750
 CC K751
 CC K752
 CC K753
 CC K754
 CC K755
 CC K756
 CC K757
 CC K758
 CC K759
 CC K760
 CC K761
 CC K762
 CC K763
 CC K764
 CC K765
 CC K766
 CC K767
 CC K768
 CC K769
 CC K770
 CC K771
 CC K772
 CC K773
 CC K774
 CC K775
 CC K776
 CC K777
 CC K778
 CC K779
 CC K780
 CC K781
 CC K782
 CC K783
 CC K784
 CC K785
 CC K786
 CC K787
 CC K788
 CC K789
 CC K790
 CC K791
 CC K792
 CC K793
 CC K794
 CC K795
 CC K796
 CC K797
 CC K798
 CC K799
 CC K800
 CC K801
 CC K802
 CC K803
 CC K804
 CC K805
 CC K806
 CC K807
 CC K808
 CC K809
 CC K810
 CC K811
 CC K812
 CC K813
 CC K814
 CC K815
 CC K816
 CC K817
 CC K818
 CC K819
 CC K820
 CC K821
 CC K822
 CC K823
 CC K824
 CC K825
 CC K826
 CC K827
 CC K828
 CC K829
 CC K830
 CC K831
 CC K832
 CC K833
 CC K834
 CC K835
 CC K836
 CC K837
 CC K838
 CC K839
 CC K840
 CC K841
 CC K842
 CC K843
 CC K844
 CC K845
 CC K846
 CC K847
 CC K848
 CC K849
 CC K850
 CC K851
 CC K852
 CC K853
 CC K854
 CC K855
 CC K856
 CC K857
 CC K858
 CC K859
 CC K860
 CC K861
 CC K862
 CC K863
 CC K864
 CC K865
 CC K866
 CC K867
 CC K868
 CC K869
 CC K870
 CC K871
 CC K872
 CC K873
 CC K874
 CC K875
 CC K876
 CC K877
 CC K878
 CC K879
 CC K880
 CC K881
 CC K882
 CC K883
 CC K884
 CC K885
 CC K886
 CC K887
 CC K888
 CC K889
 CC K890
 CC K891
 CC K892
 CC K893
 CC K894
 CC K895
 CC K896
 CC K897
 CC K898
 CC K899
 CC K900
 CC K901
 CC K902
 CC K903
 CC K904
 CC K905
 CC K906
 CC K907
 CC K908
 CC K909
 CC K910
 CC K911
 CC K912
 CC K913
 CC K914
 CC K915
 CC K916
 CC K917
 CC K918
 CC K919
 CC K920
 CC K921
 CC K922
 CC K923
 CC K924
 CC K925
 CC K926
 CC K927
 CC K928
 CC K929
 CC K930
 CC K931
 CC K932
 CC K933
 CC K934
 CC K935
 CC K936
 CC K937
 CC K938
 CC K939
 CC K940
 CC K941
 CC K942
 CC K943
 CC K944
 CC K945
 CC K946
 CC K947
 CC K948
 CC K949
 CC K950
 CC K951
 CC K952
 CC K953
 CC K954
 CC K955
 CC K956
 CC K957
 CC K958
 CC K959
 CC K960
 CC K961
 CC K962
 CC K963
 CC K964
 CC K965
 CC K966
 CC K967
 CC K968
 CC K969
 CC K970
 CC K971
 CC K972
 CC K973
 CC K974
 CC K975
 CC K976
 CC K977
 CC K978
 CC K979
 CC K980
 CC K981
 CC K982
 CC K983
 CC K984
 CC K985
 CC K986
 CC K987
 CC K988
 CC K989
 CC K990
 CC K991
 CC K992
 CC K993
 CC K994
 CC K995
 CC K996
 CC K997
 CC K998
 CC K999
 CC K1000
 CC K1001
 CC K1002
 CC K1003
 CC K1004
 CC K1005
 CC K1006
 CC K1007
 CC K1008
 CC K1009
 CC K1010
 CC K1011
 CC K1012
 CC K1013
 CC K1014
 CC K1015
 CC K1016
 CC K1017
 CC K1018
 CC K1019
 CC K1020
 CC K1021
 CC K1022
 CC K1023
 CC K1024
 CC K1025
 CC K1026
 CC K1027
 CC K1028
 CC K1029
 CC K1030
 CC K1031
 CC K1032
 CC K1033
 CC K1034
 CC K1035
 CC K1036
 CC K1037
 CC K1038
 CC K1039
 CC K1040
 CC K1041
 CC K1042
 CC K1043
 CC K1044
 CC K1045
 CC K1046
 CC K1047
 CC K1048
 CC K1049
 CC K1050
 CC K1051
 CC K1052
 CC K1053
 CC K1054
 CC K1055
 CC K1056
 CC K1057
 CC K1058
 CC K1059
 CC K1060
 CC K1061
 CC K1062
 CC K1063
 CC K1064
 CC K1065
 CC K1066
 CC K1067
 CC K1068
 CC K1069
 CC K1070
 CC K1071
 CC K1072
 CC K1073
 CC K1074
 CC K1075
 CC K1076
 CC K1077
 CC K1078
 CC K1079
 CC K1080
 CC K1081
 CC K1082
 CC K1083
 CC K1084
 CC K1085
 CC K1086
 CC K1087
 CC K1088
 CC K1089
 CC K1090
 CC K1091
 CC K1092
 CC K1093
 CC K1094
 CC K1095
 CC K1096
 CC K1097
 CC K1098
 CC K1099
 CC K1100
 CC K1101
 CC K1102
 CC K1103
 CC K1104
 CC K1105
 CC K1106
 CC K1107
 CC K1108
 CC K1109
 CC K1110
 CC K1111
 CC K1112
 CC K1113
 CC K1114
 CC K1115
 CC K1116
 CC K1117
 CC K1118
 CC K1119
 CC K1120
 CC K1121
 CC K1122
 CC K1123
 CC K1124
 CC K1125
 CC K1126
 CC K1127
 CC K1128
 CC K1129
 CC K1130
 CC K1131
 CC K1132
 CC K1133
 CC K1134
 CC K1135
 CC K1136
 CC K1137
 CC K1138
 CC K1139
 CC K1140
 CC K1141
 CC K1142
 CC K1143
 CC K1144
 CC K1145
 CC K1146
 CC K1147
 CC K1148
 CC K1149
 CC K1150
 CC K1151
 CC K1152
 CC K1153
 CC K1154
 CC K1155
 CC K1156
 CC K1157
 CC K1158
 CC K1159
 CC K1160
 CC K1161
 CC K1162
 CC K1163
 CC K1164
 CC K1165
 CC K1166
 CC K1167
 CC K1168
 CC K1169
 CC K1170
 CC K1171
 CC K1172
 CC K1173
 CC K1174
 CC K1175
 CC K1176
 CC K1177
 CC K1178
 CC K1179
 CC K1180
 CC K1181
 CC K1182
 CC K1183
 CC K1184
 CC K1185
 CC K1186
 CC K1187
 CC K1188
 CC K1189
 CC K1190
 CC K1191
 CC K1192
 CC K1193
 CC K1194
 CC K1195
 CC K1196
 CC K1197
 CC K1198
 CC K1199
 CC K1200
 CC K1201
 CC K1202
 CC K1203
 CC K1204
 CC K1205
 CC K1206
 CC K1207
 CC K1208
 CC K1209
 CC K1210
 CC K1211
 CC K1212
 CC K1213
 CC K1214
 CC K1215
 CC K1216
 CC K1217
 CC K1218
 CC K1219
 CC K1220
 CC K1221
 CC K1222
 CC K1223
 CC K1224
 CC K1225
 CC K1226
 CC K1227
 CC K1228
 CC K1229
 CC K1230
 CC K1231
 CC K1232
 CC K1233
 CC K1234
 CC K1235
 CC K1236
 CC K1237
 CC K1238
 CC K1


```

Db 709 LRKLKYLGSQVGFVTKGKWMIPGSESIKIPVCIKIVTEDSGRSQFQAVTDHMLTIGSLDH 768
Qy 780 PIVSRLIGICLSTVOLVQMLPYGCLLDHVRNRRGLSQODILNMQIAKMSYLEDV 839
Db 769 AHIVRLIGICPSSSLQLVQYGLPLSLDHDVHQHRLGLAQQLLLNMGVQIAKMSYLEEH 828
Qy 840 RLVRHRLAARNTLVKSPNHYKIDPGLARLLDDIDFHYHADGKGVPIKMMALLESILRRR 859
Db 829 GNVHRLAARNTLVKSPQVADLPDDLDQLVSEAKTPKMMALLESILRRR 868
Qy 900 THQSDVMSGVTVWMLMTFEGAKPYDGIIPAREIPDLSEKGERLPQPPICITDVMYMKW 959
Db 889 THQSDVMSGVTVWMLMTFEGAKPYDGIIPAREIPDLSEKGERLPQPPICITDVMYMKW 948
Qy 960 MIDSECRPRPRVSEPSRMAQDQRFVVIQNEIDLPA---SPLDSTFYRSLLEDDMDG 1016
Db 949 MIDENRPTFKELANFTFARMADPPRYLVIKRES-GRGIAPGEPRHGLTNKKLEVELE 1007
Qy 1017 LVDAEYLVPCQGFPCPDPAFGAGGVHHRHSSSTRSGGDIITGLER-SEEARSRPL 1075
Db 1008 ELDDLIDLLEAEED-NTATTTLSALSLPVGTLNRRPGRSQSL 1048
Qy 1076 APSEGAQDVFPDGLGMAKAGLSLPTHD-PSPLQRYSEDPTVPLP-----SETDGYV 1128
Db 1049 SSSSGI-MEMNGCNLSESCQESAVSGSSERCPVSLH-----PWRGCLASESSEGHV 1101
Qy 1129 A-----PLTCSQDP-YNQPPVPRPSPRSGP-----L 1157
Db 1102 TQSEAELOEKVSWCRSRNSRSPRPGRDASVHSGHSLLPVPLSPGLEEDVNGYV 1161
Qy 1158 PAARPGATLERKXTISP-GKNV-----KVPRFAGAVENPEVLTTCGGAAPQHP 1210
Db 1162 PTHLNGTPSSRSGTSLVGLSVLTGTEBED-EEYEVNRRRRRSP-PHP 1212
Qy 1211 PAFSPAFLNLYVD-----QDPERGAPSTFGTPTANPEYL 1249
Db 1213 RRSSELELGYVDGSDLSASLGTSQCPHVPVMTPTGTPDEDEYEM 1263

RESULT 10
ERB3 RAT STANDARD; PRT; 1339 AA.
AC 062759; 062955;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-3 precursor (BC 2.7.1.112)
DE (c-erbB3).
GN ERB3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RC MEDLINE=96096535; PubMed=8522190;
RA Hellyer N.J., Kim H.-H., Greaves C.H., Sierke S.L., Koland J.G.;
RT "Cloning of the rat ErbB3 cDNA and characterization of the
RT recombinant protein."
RL Gene 165:279-284 (1995).
RN [2]
RP REVISIONS TO 85; 513 AND 565.
RA Hellyer N.J., Koland J.G.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 922-1097 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Sciatic nerve;
RC MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Froehner P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuroligin and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration."
RL J. Neurosci. 17:1642-1659 (1997).

```

```

CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGUINS AND NTRK.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
CC PHOSPHATIDYLINOSITOL 3-KINASE.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U29339; AAC28498.2; -
DR EMBL; U52530; AAC53050.1; -
DR HSSP; P13352; 1FGK.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR000719; Prot Kinase.
DR InterPro; IPR001245; Tyr kinase.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR0109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 5.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Trasnembi; Glycoprotein; Multigene family; Receptor; Signal;
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
KM
FT SIGNAL 1 19
FT CHAIN 20 1339
FT DOMAIN 20 643
FT TRANSMEM 644 662
FT DOMAIN 663 1339
FT DOMAIN 183 259
FT DOMAIN 707 964
FT NP_BIND 713 721
FT BINDING 740 740
FT ACT_SITE 832 832
FT DISULFID 186 194
FT DISULFID 190 202
FT DISULFID 210 218
FT DISULFID 214 226
FT DISULFID 227 235
FT DISULFID 231 243
FT DISULFID 246 255
FT DISULFID 259 286
FT DISULFID 290 301
FT DISULFID 305 320
FT DISULFID 323 327
FT DISULFID 500 509
FT DISULFID 504 517
FT DISULFID 520 529
FT DISULFID 532 549
FT DISULFID 536 573
FT DISULFID 576 585
FT DISULFID 589 610
FT DISULFID 613 621
FT DISULFID 617 629
FT CARBOHYD 126 126
FT CARBOHYD 250 250

```

```

PT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 566 566 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CONFLICT 1028 1028 L -> P (IN REF. 3).
SQ SEQUENCE 1339 AA; 147545 MW; 0AA5F2402BBDFE1E CRC64;

Query Match 34.84; Score 2370.5; DB 1; Length 1339;
Best Local Similarity 41.18; Pred. No. 6,7e-118;
Matches 527; Conservative 170; Mismatches 431; Indels 155; Gaps 34;

3 LALCRGMLLALLPSPA---STQVCTGDMKRLRLASPEITLMDRLHYOGQOVQGN 59
7 LOYLCC---FLSLARSGEKGNGQAVCPGTLNGLSTVDADNOYQITLYKLYEKCEVVMGN 62
60 LELTYLPNTASLSPLDIOEVOGYVLIHNOVRQVPLQRLRYRGTLFEDNYALAVLDN 119
63 LEIVLGNHNDLSFLQWIREVTGYLVANMEFVLPNLRYVARGTQVDDGKFAIFVM-- 120
120 GDEPLANTPTTGASPGSLRLQLRLSTLTKGCVLIGRNPOLCYQDTITLKDTFHKNNOL 179
121 ---LNNNT---NSSHALRLQKFTOLTEILSGVYIEKNDKLCMDITIDRDIVVR-- 170
180 ALTLITNRSRACHGSPMCKGSRCKMGESSEDCSLTRTVACAGGC-ARCKPLPTDCHE 238
171 GAETVANKANGANCPCHIEVCKG-RCMGPGDDCQILTKITCAPQCNCRCGPPNQCCHD 229
239 QCAAGCTGPRHSDCLACLAHFNHSGICELHCPALVTYNTDFESNPBEGRYTFGASCTA 298
230 ECAGGCGSGPDPTCPACRRFNDGACVRCBPBLVYKTLFQLEPNHTKYQGVGVAS 289
299 CPTNYLSTVGSCTLVCPRLNNOETADGTORCKSKCPACVYCGIMENLEKAVYTS 338
290 CPHNPFV-VQDTFCVRACPPDKMEVD-RHGLKMEPCGGLPKCKCEGSGG-SRYQTVD 345
359 ANIOEPAGCKKIFGSLAFLEPESFGDPASNTAPLOPQLOVPEITLITGYLYISAPDS 418
346 SNIDGPNCKIKIGNDLFTGLANDVDMKIPALDEPKLVNFTVRITITGLYINQSPPH 405
419 LPLDSVFONLQVIRGRILHNGAVS-LTLQGLKISWGLARSLRELKSGALVHHNTILCFV 477
406 MHNFSVFNLTITIGRSLVNRGSLIMKMLNTSLGFRSLKETSARAVYISANQQLCYH 465
478 HTVPMDQLFNPQIALHTA-NRPEDECVGBGLACHOLCARHGWGEPPTQCVCNCSQFLR 536
466 HSLMTTRLLRGPSEBRRLDIKYDRPLGELAGKVCDDLSSGGCGWGPQCLSCRYYSR 525
537 GQECVECEKVLQGLPREYVNAHGLPGHPCQOPNGSVTCGPEADOCVACAHYKDDPFC 596
526 EGVCAVTHCNFLQSBPFBVBAQFSCAPELPMEGSTGCGSSGSDACARCAHRRDQPHC 585
597 VAPCGSVKEDLSYMPIMKFPDEEGACQCPINTHSC--VDLDDKCPABORASPLTSI 654
586 VNGCPHGLG-AKGPYKYPDAQNECRPHENCTOGCNPBLDCLGQAEVLMKSHLV 643
655 ISAVGILVAVLVGVFGLIKRQOKIR-KYTRRLLOETELVEPLTPSGAMENQAKMR 713
644 IAVTVG--LAVIIMLIGSGFLYMRGRIRQNRARRYLGESESTIEPLDPG-EKANKYLAR 700
714 ILKSTELRKVYLSGSGAGTVYVGIWIPDBENVYKIPALVLRNENTSPRANKELIDBAV 773
701 IFKSTELRKVYLSGSGAGTVYVGIWIPDBENVYKIPALVLRNENTSPRANKELIDBAV 760
774 MAGVGSFYVRLGLICLTSTVQVLTQAMPYGLLDHRENRGRGLSDLLNMQWQAKGM 833
761 VGSIDHNIYRLGLICPSSSLQVLTQVYLPGLSLDHKQKRETIQPOLLMNQWQAKGM 820
834 SYLEDVRLVHRDLAARNVLVKSPPNHYKITDFGLARLLDIDETEHADGKVPIMMALES 893

```

```

DB 821 YVIEHSWHRDLALANNMLKSPSQVADFGVADLLPDDKOLLHSEAKTPIKMALES 880
QY 1LRRTTHSDVMSYGVTVWELMTFGAKPYDGIPIAEIPDLLEKGERLPORPICTIDVYM 953
DB 881 IHFGKTHSDVMSYGVTVWELMTFGAKPYDGIPIAEIPDLLEKGERLPORPICTIDVYM 940
QY 954 IMVCKMIDSECPRRRELVSFSPRNARDPQRFVIONEDLPASPLDSTFPYRSLBEDD 1013
DB 941 VVYCKMIDENIPTEKELANETFRMARDPRLVIRKAS-GEQTP--PAAPSVLTATKE 997
QY 1014 MEDLVAAEITVQOGFCDDPARAGAGVHHHRSSSTSSGGDLTLGLEPSE----- 1068
DB 998 L-----QEALBELL-----DPLDLEAEEGLATS 1023
QY 1069 -----BAPSPPLAPSEG-----AGSDVPDGLGAKAGLQSLPTHD 1105
DB 1024 LGSALSLPTGTLTPRSGSGLSPSSGYMPMNGSSIGELADBAVLGGRQFSPISLH- 1082
QY 1106 PSPLQRYSEDPVPLPSETDGYV-----APL-----TC-----SPOPE-----YVNOPOV 1145
DB 1083 PIPRGR-----PASESSEGHVTGSEAELOEKVYCSRSRSRSPRPGDSAVHSORHS 1135
QY 1146 RPOPSPRECP-----LPARPGATILERPPTTSP-GNGGV-----KDYFAF 1167
DB 1136 LTPVTPSPGLEBEDNGYVMPDTHLRKASSRSGTSSVGLSVLGTBEEDD----- 1191
QY 1188 GSAVENPEYLLTPOGGAPOPHPP 1210
DB 1192 -----EEVETMNRKRGRSP-PRPP, 1209

RESULT 11
EGFR DROME STANDARD; PRT: 1426 AA.
ID P0412; O18370; 061601; P81668; Q9WZG0;
AC 13-AUG-1987 (Rel. 05, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (EGfr)
DE (Garten receptor) (Tropo protein) (Drosophila relative of ERBB).
GN EGFR OR TOP OR C-ERBB OR DER OR CSJ10079.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid:7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS TYPES I AND II).
RX MEDLINE=94350209; PubMed=8070664;
RA Clifford R., Schnupbach T.;
RT "Molecular analysis of the Drosophila EGF receptor homolog reveals
RT that several genetically defined classes of alleles cluster in
RT subdomains of the receptor protein.";
RL Genetics 137:531-550 (1994).
RN [2]
RP REVISIONS.
RA Clifford R., Schnupbach T.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=65124611; PubMed=2982499;
RA Livneh E., Glazer L., Segal D., Schlessinger J., Shilo B.-Z.;
RT "The Drosophila EGF receptor gene homolog: conservation of both
RT hormone binding and kinase domains.";
RL Cell 40:599-607 (1985).
RN [4]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RX STRAIN=Oregon-R; TISSUE=Embryo;
RX MEDLINE=87002474; PubMed=3093080;
RA Schejter E.D., Segal D., Glazer L., Shilo B.-Z.;
RT "Alternative 5' exons and tissue-specific expression of the
RT Drosophila EGF receptor homolog transcripts.";
RL Cell 46:1091-1101 (1986).

```

[5] SEQUENCE FROM N.A. (ISOFORMS TYPE I AND TYPE II), TISSUE SPECIFICITY, AND MUTATION ANALYSIS.
 RX MEDLINE=99102120; PubMed=9882502;
 RA Leoshkin A.M., Yu S.-Y., Katz J., Baker N.E.;
 RT "Several levels of EGF receptor signaling during photoreceptor specification in wild-type, Ellipse, and null mutant *Drosophila*.";
 RL Dev. Biol. 205:129-144(1999).
 [6]
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I).
 RC STRAIN=Beckley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Adiyanti A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bence P.V., Bereman B.P., Bhandari D., Bolhakov S.,
 RA Borovica D., Botchan M.R., Bouck J., Brocktein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mayhew A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doult L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durlin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foutel C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodex C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibeagwa C.,
 RA Jaimali M., Kalush P., Karpis G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laebo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mates J.B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshire A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palczolzi M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen T.,
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
 RA Spletter R., Spretling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitkies R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [7]
 RP SEQUENCE OF 959-1078 FROM N.A.
 RC STRAIN=Daekwanmyeong;
 RX MEDLINE=85137938; PubMed=2983232;
 RA Wadsworth S.C., Vincent W.S. III, Bildeau-Wentworth D.;
 RT "A *Drosophila* genomic sequence with homology to human epidermal growth factor receptor.";
 RL Nature 314:178-180(1985).
 [8]
 RP SEQUENCE OF 1069-1121 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=98401146; PubMed=9731193;
 RA Oates A.C., Wollberg P., Achen M.G., Wilke A.F.;
 RT "Sampling the genomic pool of protein tyrosine kinase genes using the polymerase chain reaction with genomic DNA.";
 RL Biochem. Biophys. Res. Commun. 249:660-667(1998).
 [9]
 RP SEQUENCE OF 1133-1137, 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION ANALYSIS.
 RX MEDLINE=92038942; PubMed=1936959;
 RA Raz E., Schejter E.D., Shilo B.Z.;
 RT "Interallelic complementation among DER/fib alleles: implications for the mechanism of signal transduction by receptor-tyrosine kinases.";

RL Genetics 129:191-201(1991).
 RN [10]
 RP REVIEW.
 RX MEDLINE=97248481; PubMed=9094709;
 RA Perlmutter N., Perkins L.A.;
 RT "There must be 50 ways to rule the signal: the case of the *Drosophila* EGF receptor.";
 RL Cell 89:113-116(1997).
 CC -1- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS, WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAF-MAPK PATHWAY. INVOLVED IN A MYRIAD OF DEVELOPMENTAL DECISIONS.
 CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES AND FOR THE DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL POLARITIES OF THE OOCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF ANTIOSEROSA AND VENTRAL NEUROECODERMAL CELLS, GERM BAND RETRACTION, CELL FATE SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF CUTICLE.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Type I;
 CC IsoId=P04412-1; Sequence=Displayed;
 CC Name=Type II;
 CC IsoId=P04412-2; Sequence=VSP_002897;
 CC -1- TISSUE SPECIFICITY: UBIGUOUSLY EXPRESSED IN EMBRYOS. IN LARVAE, UNIFORM EXPRESSION IS SEEN IN WING DISKS, GENITAL DISK, ANLAGEN OF TESTIS AND OVARY AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW, LEVELS REMAIN HIGH IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX AND THORACIC AND ABDOMINAL GANGLIA.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AF053754; AAC08536.1; -;
 DR EMBL, AF053753; AAC08536.1; JOINED.
 DR EMBL, AF053754; AAC08535.1; -;
 DR EMBL, AF053752; AAC08535.1; JOINED.
 DR EMBL, K03054; AAAS1462.1; -;
 DR EMBL, K03417; AAAS1460.1; -;
 DR EMBL, K03416; AAAS0965.1; -;
 DR EMBL, AF109077; AAD26134.1; -;
 DR EMBL, AF109076; AAD26132.1; -;
 DR EMBL, AF109082; AAD26132.1; JOINED.
 DR EMBL, AF109078; AAD26133.1; -;
 DR EMBL, AF109084; AAD26133.1; JOINED.
 DR EMBL, AF109079; AAD26130.1; -;
 DR EMBL, AF109081; AAD26130.1; JOINED.
 DR EMBL, AF109079; AAD26131.1; -;
 DR EMBL, AF109083; AAD26131.1; JOINED.
 DR EMBL, AF109080; AAD26135.1; -;
 DR EMBL, AF003454; AAF46732.1; -;
 DR EMBL, X02293; CAA26157.1; -;
 DR EMBL, AF003912; CAA05747.1; -;
 DR EMBL, X78930; CAA55523.1; -;
 DR EMBL, X78918; CAA55521.1; -;
 DR EMBL, X78919; CAA55522.1; -;
 DR HSSP, P11362; 1FGK.
 DR FlyBase: FBgn0003731; Egfr.
 DR GO: GO:0007469; P:antennal morphogenesis; NAS.
 DR GO: GO:0003481; P:eggshell pattern formation; IGI.
 DR GO: GO:0007390; P:germ-band shortening; IMP.
 DR GO: GO:0007444; P:imaginal disc development; IMP.

DR GO:0007479; P-leg disc proximal/distal pattern formation; IMP.
 DR GO:0008671; P-maternal determination of dorsal/ventral ax. . .; IMP.
 DR GO:0007477; Pinotum morphogenesis; IMP.
 DR GO:0007314; Pinocyte anterior/posterior axis determination; NAS.
 DR GO:0045468; P-regulation of R8 spacing; NAS.
 DR GO:0016330; P-second mitotic wave (sensu Drosophila); IMP.
 DR GO:0007476; P-wing morphogenesis; IMP.
 DR InterPro: IPR000494; Egfr L domain.
 DR InterPro: IPR006211; Furin-like.
 DR InterPro: IPR006212; Furin repeat.
 DR InterPro: IPR000719; Prot. Kinase.
 DR InterPro: IPR001245; Tyr. Kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.

Query Match 29.0%; Score 1975; DB 1; Length 1426;
 Best Local Similarity 33.0%; Pred. No. 5,7e-97;
 Matches 470; Conservative 164; Mismatches 432; Indels 338; Gaps 41;

24 QVCTGDMKRLRLPASPTLMDLRHLHYGCGVVGSLTYLPT-NASLFLDIDIOVG 82
 100 KICIGTKRSLSVPSNKHNRNLRDRYTNCTYVDGNLKTWLPNENLDSFLDNIREVTG 159
 83 YVLIANOVROVPLQRLIRYGTQLP-----EDNYALAVLDNGDPLANTTPVTSAGSG 137
 160 YLIIHVADVKVVPPLQIIRGRITLSVSEERKALFV-----TYSKM 203
 138 RELQLSLTEILKGVLIQNPOLCYODTILMKDIFKNNQLALLIDTNRSPACHPSP 197
 204 YLLEIIDLADLVNGVGFNNVNLCHMTITQNSEIVNGTDAYYNDFTAPRECEPCHE 263
 198 MCKGSHCWSESSDCSLTRTVAGGA--RCKGRLPTDCCHGCAAGCTGPRHSDCLAC 255
 264 SCTHG-CWEGSPKXCKPSKLTCSPOCAGRCYGPKEKRECHLFCAGGCTGPTOKDIAC 322
 256 LPHNSGICELHPALVTYTDTFESMPNREGRYTPGASCTVACPXYLSTVDGSCVLVC 315
 323 KPFDEAVSSEPCPRKKNPTTYVLETNPEGKAYAGATCVKSCP--GHLRLDNGACVRS 361
 316 PLHNOEVTABDGTORCEKSCPCARVYGLGMEHLREVRVANSANTIOEPAGCKKIFGSLA 375
 382 PDDKMDKGE-----CVRNGPCPKTCPTGVTVLH-----AGNIDSPKCTVIDENIR 428
 376 FLPESEDG--DPASNTA-----PLOBOLQVEFLEBITGYLYISAMPDPLDLSVFON 427
 429 ILDTGSGGQDVYANTWGPRIYLPDBERBVSFTVSEIIGYINLSETHQFNLISFRN 468
 428 LOVIRIRILHNGAY-SITLGLGISWLGSLRELSSGLALLIHNTHLCVHTVPMDF 466
 489 LETIHROLMESMFALALAYKSSLYSLEMENLKQISSGSAVIOHNRDLCTVSNIRMPAIQ 548
 487 RNPQALLHTANREDECVGEGLAQHQLCARHGMWGSPGPOCNCSQFLAGSCVBEKRV 546
 549 KEPEQVWVWENIRADLCENKGTICSDQCNEDCGWAGTDQCLTKNFENGTCLADCGY 608
 547 LOGLPREYVNAHCLPCHREPCOPNGSVTCFGEADQCVACAHYKDPFCVACAP----- 601
 609 ISNAAYK--PDNRCTCKIHPRECR-----TCNGAGADHCQGCYVVRDQCHCVSECPKKN 660
 602 -----SGVK-----CPI-----NCH-----PLSTWPIW 614
 661 DRGVCRCHATCDGCTGPKDTIGIGACTTCNLAIINNDATVKECLDKDCPD--GY--FW 717
 615 KF--PREBAGACP-----CPI-----NCH-----632
 718 EYVHPQGSLLKLAGRAVCRKCHPLCELTNYGHEOVSCKTHYKRRQCECEPADH 777
 633 -----SC-----VLDKDG-----CPAQ 646
 778 YNDBEQRECGQRHRECGCTGPADDCSGRNPLPLFANETGPVNSTMNCSTKCLEM 837
 647 R-----ASPLTS-----IISAVVGILLVVVLGVVFGILLKRRQ 679

DB 838 RHYNYOTAIIPYCAASPPRSSKITANLDVNMFIITGAVLVPTICLCV--TYICRQK 895
 QY 680 QKIRKKT--MRLLQETELVEPLTPSGAMPNCAOMRLKETELRKVYLSGAGFYVYG 737
 DB 896 QKAKKETVMTALSCGDBSEPLRPISNIGANLCKLRIVKQAEIRKSGVLGAGGRVYKG 955
 QY 738 IWPDSGNKIPPAIVLRENTSPKANKETLDAAYVAGSPVSVLLGLTSTVQLV 797
 DB 956 VWVPEBENYKIPPAIKELKSTGASSESEPLREAYVMASEHVLLKLAVCMSSOMLI 1015
 QY 798 TOLMPYGLLDHYRENRGLSGODLNMWCMQIKGMSYLEDVRLVHRDLAARVLYVSPN 857
 DB 1016 TOLMPYGLLDHYRENRGLSGODLNMWCMQIKGMSYLEDVRLVHRDLAARVLYVSPN 1075
 QY 858 HVKIDTFGALRLDIDETEHADGKVPKIMMALLESILRRRFTQSDVMSYGVTVWELMT 917
 DB 1076 LVKITPFGALKLSSDSNNEYKAGGMPIKMLLBICIRNVFYSKDPVMAFGVIMWELT 1135
 QY 918 FGAKPYDGIIPAREIPDLLEKGERLPPOPCTIDVYIMVCKMIDSECRPRELVSBSFS 977
 DB 1136 FGORPHENIPAKDIPDLIEVGLLEQBPICSLDIYCTLLSCWHLDAAMRPTFKOLTVFA 1195
 QY 978 RMAARDQRFVYVQNEBLS--PASPLDSTFYSLLED--DMGDLVDAEYLVPOQGFPC 1032
 DB 1196 EPARDPGRYALPGDFRPLA-----TTSODEKDLIKLAPTTGSEALAKPDDYLQ 1248
 QY 1033 PDPAPGAGGVHRRHSSSTRSGCDLTLGLEPSEBEAP-----RSPLAPSEGAGSDVF 1086
 DB 1249 EKAPGPS-----HTDCT-----DEMPLKLNKYCKDPSKNSSTGDER 1267
 QY 1087 DG--DLGMAKAGLOSPLTHDPSPLORYSEDPVLPSTGTVAPLTCSPQPEYVNO 1143
 DB 1288 DSGAREVGQNLK-----LDLPVDEDDYLMF--TCOPGNNNNNM 1325
 QY 1144 DVPAPPSRPEGLPARAPGATLERPKTLPCKNGVVKOVPAFGAVENPEYL-----TP 1199
 DB 1326 N-----NPNQNNMAVGAAGYM-----DLIGVPVSDNPEYLLAQT 1364
 QY 1200 QGGAAPQPH-----PPAPSP--AFDNLVYMD 1224
 DB 1365 GYVESPIPTOTIGIPWGGRGTMKAVVPMGSEPTSSDHYND 1408

RESULT 12
 ERBB ALV STANDARD; PRT; 634 AA.
 AC P00534;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase transforming protein erbB (EC 2.7.1.112).
 GN V-ERBB.
 OS Avian leukosis virus.
 OC Viruses; Retroviral viruses; Retroviridae; Alpharetrovirus.
 OC NCBI TaxID=11864;
 RX MEDLINE=85228222; PubMed=2988784;
 RP SEQUENCE FROM N.A.
 RA Nilsen T.W., Maroney P.A., Goodwin R.G., Rotman F.M.,
 RA Cittenberg U.B., Raines M.A., Kung H.-J.,
 RT "c-erbB activation in ALV-induced erythroblastosis: novel RNA
 RT processing and promoter insertion result in expression of an
 RT amino-truncated EGF receptor";
 RL Cell 41:719-726 (1985).
 CC -|- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -|- MISCELLANEOUS: THIS PROTEIN IS PRODUCED BY ERYTHROLEUKEMIA CELLS
 CC IN CHICKS WHICH HAVE BEEN INJECTED WITH THE AVIAN LEUKOSIS VIRUS
 CC AND SUBSEQUENTLY DEVELOP LEUKEMIA. THE LEUKEMIA ARE INDUCED BY
 CC THE INSERTION OF RETROVIRAL PROMOTER ELEMENTS IN OR NEAR THE
 CC C-ERBB GENE, WHICH LEADS TO C-ERBB ACTIVATION.
 CC -|- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ENV-ERBB
 CC PROTEIN.

CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC EMBL: M10066; AAA48763.1; ALT_INIT.
 DR HSSP: P11362; 1RGK.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase.1.
 DR SMART: SM00219; TyrcKc_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR TRANSFERASE; Tyrosine-protein kinase; ATP-binding; Oncogene;
 KM Glycoprotein; Phosphorylation.
 FT DOMAIN 132 399 PROTEIN KINASE.
 FT NP_BIND 138 146 ATP (BY SIMILARITY).
 FT BINDING 165 165 ATP (BY SIMILARITY).
 FT ACT_SITE 257 257 BY SIMILARITY.
 SQ SEQUENCE 634 AA; 70891 MW; E705E33A0B01PCC CRC64;
 Query Match 25.7%; Score 1748.5; DB 1; Length 634;
 Best Local Similarity 52.3%; Pred. No. 2.1e-85;
 Matches 370; Conservative 79; Mismatches 135; Indels 123; Gaps 17;
 QY 587 CAHYKDPFCVACPSGVKPDLSYMPKRPDESGACOPCINCTHSCVLDLDDGCGAEG 646
 DB 3 CAHFDGPHCVKACPSVGLSNDTL-VKXADANAVQCLCPNCTRGCKGPGLEGCP--- 58
 QY 647 RASPLTISIAVY-GILLVVVGVVGFILIKRQOKIRKYMRLLOETLEVEPLPSGA 705
 DB 59 NGSKTPIAGVGVGLCLVVGIGLILYRR-HYRKQTLRLLOERLEVEPLTSGE 117
 QY 706 MPMQAKRIKETELKRVKVLGSGAFYTKGIPIPGENVKIPVAIKLRENTSPANK 765
 DB 118 APQAKRLIKETEFKKVKVLGSGAFYTKGIPIPGENVKIPVAIKLRENTSPANK 177
 QY 766 EILDEAVYVAGVSPVYSLGICLTSTVOLVQIMPYGCLDHRNRRLSGODLLN 825
 DB 178 EILDEAVYVAGVSPVYSLGICLTSTVOLVQIMPYGCLDHRNRRLSGODLLN 237
 QY 826 CMOIAGKMSYLEVRLVHRDLAARVYKSPNHVKTDPGLARLLDDETEYHADGKVP 885
 DB 238 CVOIAGKMSYLEVRLVHRDLAARVYKSPNHVKTDPGLARLLDDETEYHADGKVP 297
 QY 886 IKMALESILRRPFTHOSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPP 945
 DB 298 IKMALESILRRPFTHOSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPP 357
 QY 946 ICTIDVYVIMVCKMIDSECRPRRELVSERWAPRQRFVVO-NEDGAPASPLDSTF 1004
 DB 358 ICTIDVYVIMVCKMIDSECRPRRELVSERWAPRQRFVVO-NEDGAPASPLDSTF 417
 QY 1005 YRSLLDDDDMDLVDAEYLVPOGFCPPDPAFGAGVHHRRHSSSTRSGGGLTLGLE 1064
 DB 418 YRSLLDDDDMDLVDAEYLVPOGFCPPDPAFGAGVHHRRHSSSTRSGGGLTLGLE 449
 QY 1065 PSEEEAPRSFL-----APSGAGSDVFDGLGMAKAGLSLPHDPSPLQRYSEDPVTP 1119
 DB 450 -----SRTEPLSLSLATSNNSATNCID-----NNGQGHFVREDSFVQRYSDPTGN 495
 QY 1120 LPSET--DGVAVPLTCSPOPEVYNQPRQPRSPRSGPLPAARPAATLERPITLSPGK 1177
 DB 496 FLBSRIDGFL-----PAPRYVNO--LMPKPS-----TAMVQ 526

QY 1178 NGVVKDF-----AFGAVENPEYLTPGGAAPQPHPPAFSPAFDNLVY 1222
 DB 527 NGVYNNISLTAISKLPMDSRQNSHSTAVDNPETL-----NTNOSPLAKTVESSPY 578
 QY 1223 WDC-----DPPE-----RGAPSTKGPPTLNPENYLGIDVP 1254
 DB 579 WISGNGHQLNDNDYQODFLPNETKENGLIKVPAENPEYLRYAAP 625
 RESULT 13
 ERBB_AVIER STANDARD; PRT; 604 AA.
 ID ERBB_AVIER
 AC P00353;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase transforming protein erbb (BC 2.7.1.112).
 GN V-ERBB.
 OS Avian erythroblastosis virus (strain BS4).
 OC Viruses; Retroviruses; Retroviridae; Avian type C retroviruses.
 OC NCBI_TaxId=79685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H;
 RX MEDLINE=84026539; PubMed=6313229;
 RA Yamamoto T., Nishida T., Miyajima N., Kawai S., Ooi T., Toyoshima K.;
 RT "The erbb gene of avian erythroblastosis virus is a member of the src
 RT gene family";
 RL Cell 35:71-78(1983).
 RN [2]
 RP SEQUENCE OF 1-152 FROM N.A.
 RX MEDLINE=84223957; PubMed=6328658;
 RA Debutre B., Henry C., Benaisse M., Biserte G., Clavierie J.-M.,
 RA Saule S., Martin P., Stehelin D.;
 RT "Sequencing the erbb gene of avian erythroblastosis virus reveals a
 RT new type of oncogene";
 RL Science 224:1456-1459(1984).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- DISEASE: THE V-ERBB ONCOGENE TRANSFORMS AVIAN FIBROBLASTS AND
 CC ERYTHROBLASTS IN CULTURE AND INDUCES SARCOMAS AND ERYTHROLEUKEMIAS
 CC IN CHICKENS.
 CC -1- MISCELLANEOUS: V-ERBB IS A TRUNCATED AND MUTATED VERSION OF THE
 CC RECEPTOR FOR EPIDERMAL GROWTH FACTOR.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC EMBL: X02006; AAA42394.1; ALT_INIT.
 DR EMBL: X01216; AAA42400.1; --
 DR PIR: A06644; TVYU.
 DR HSSP: P11362; 1RGK.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase.1.
 DR SMART: SM00219; TyrcKc_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR TRANSFERASE; Tyrosine-protein kinase; ATP-binding; Oncogene;
 KM Glycoprotein; Phosphorylation.
 FT DOMAIN 132 399 PROTEIN KINASE.
 FT NP_BIND 138 146 ATP (BY SIMILARITY).
 FT BINDING 165 165 ATP (BY SIMILARITY).
 FT ACT_SITE 257 257 BY SIMILARITY.

FT CONFLICT 29 29 R -> W (IN REF. 2).
 FT CONFLICT 140 140 S -> F (IN REF. 2).
 FT CONFLICT 146 146 I -> V (IN REF. 2).
 SQ SEQUENCE 604 AA; 67633 MW; 76BCDD06745D609 CRC64;

Query Match 25.0%; Score 1702; DB 1; Length 604;
 Best Local Similarity 52.2%; Pred. No. 5.6e-83;
 Matches 360; Conservative 76; Mismatches 128; Indels 126; Gaps 16;

587 CAHYKPPFCVAPSGVKGKDLSTYMPKFPDEBACQPCPCINCTHSCVDLDDKCPAEQ 646
 3 CAHFIDGPHCVKACPAVGLBNDTL-VKXADANAVQLCHPNCTRCCKPGLBGC- 58
 647 RASPLTISIAVY-GILVVVVLGVGFGILIKRQOKIKXTMRLLOETELVEPLTPSGA 705
 59 NSGKTPSIAGVGVGLCLVAVGIGLGLYLRH-HIVRKQTLRLLOERELVEPLTPSGE 117
 706 MPQAOQRILKETELRKVKVLGSGAGCTVYKGIWIPGENVKIPVALKVLRENTSPKANK 765
 118 APQOAHRLIKETEFKVKVKGAGFTYKGLMIPGENVKIPVALKELRENTSPKANK 177
 766 EILDEAYVAVGSPVYSRLGLCLTSTVOLVQLMPYGLLDHVRNRRGLSODLLNW 825
 178 EILDEAYVAVSVNPHVCRLLGICLSTVOLITQLMPYGLLDYIRHKONIGSYLLNW 237
 826 CMQIAGKSTLEDRVLVHRDLAARNVLYKSPNHVKTIDFGIARLLDDETYHADGKVP 885
 238 CVOIAGKMYLERRLVHRDLAARNVLYKTPQHKITIDFGIARLLDDETYHADGKVP 297
 886 IKMALESILARRFTQSDVWSYGVTWELMTGAKYDGIPIAREIDLEKGRLLPOPP 945
 298 IKMALESILARRFTQSDVWSYGVTWELMTGSKYDGIPIAREIDLEKGRLLPOPP 357
 946 ICTIDVYMIWVKCMWIDSECRPRELVSEFSRNARDPQRFVYIQ-NEDGAPASPLDSTF 1004
 358 ICTIDVYMIWVKCMWIDSECRPRELVSEFSRNARDPQRFVYIQ-NEDGAPASPLDSTF 417
 1005 YRSLEDDDMGDLVDAEYLVPOGGFPCDPAPGAGMVHRRHSSSTRSGGDLTLGLE 1064
 418 YRLMEEBDMEDIVDAEYLVPHQGF-----NSPST----- 449
 1065 PSEEBAPRSPGL-----APSEGASDVDPDGLGMAKGLSLPHDSDPQRYSEDPTVP 1119
 450 -----SRTPVLSLSLSTNNASATNCID-----RNGQGHPRKSDSFQORYSSDPTGN 495
 1120 LPESET--DGTVAPLTCSPQPEYVNPQDPVROPSPRSGPLPAARPAATLERPRTLSPGK 1177
 496 FLEBSJDDGFL-----PABEYVNO--LMPKXSTAM----- 524
 1178 NGVVKDVFAP-----OGAVENPEYLTPOGGAPOPPPAFPAFDP 1218
 525 --VQNDIYVNIISLTATSKLPMDSRYQNSHSTAVDNPEYLT-----NTQSPKAKTVFE 574
 1219 NLYTPQDPPPERGAPSTFKGTAAENPEY 1248
 575 SSEYWIQSGNHQ-----INTLDNPDY 594

RESULT 14
 ERBB AVIEU
 ID ERBB AVIEU STANDARD; PRT; 540 AA.
 AC P11273;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase transforming protein erbb (EC 2.7.1.112).
 GN v-ERBB.
 OS Avian erythroblastosis virus (strain ts167).
 OC Viruses; Retroid viruses; Retroviridae; Avian type C retroviruses.
 NCBI_TaxID=103898;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6706458; PubMed=2876364;

RA Chol O.R., Trainor C., Graf T., Beug H., Engel J.D.;
 RT "A single amino acid substitution in v-erbB confers a thermolabile
 RT phenotype to ts167 avian erythroblastosis virus-transformed erythrocyte
 RT cells."
 RL Mol. Cell. Biol. 6:1751-1759(1986).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or
 CC send an email to license@ebi.ac.uk).

CC
 CC EMBL, M13179; AAA2401.1; --
 CC EIR; A25231; TVFEB.
 CC HSP; P11362; IFGK.
 CC InterPro; IPR000719; Prot kinase.
 CC InterPro; IPR001245; Tyr kinase.
 CC Pfam; PF00069; kinase_1.
 CC PRINTS; PR00109; TYRKINASE.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 CC PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Oncogene;
 CC Glycoprotein; Phosphorylation.
 CC DOMAIN 132 399
 CC NP BIND 138 146 ATP (BY SIMILARITY).
 CC FT BINDING 165 165 ATP (BY SIMILARITY).
 CC FT ACT SITE 257 257 BY SIMILARITY.
 CC FT VARIANT 270 270 H -> D (IN THERMOLABILE V-ERBB).
 CC SQ SEQUENCE 540 AA; 60412 MW; 5B53297AA068B65D CRC64;

Query Match 23.9%; Score 1628; DB 1; Length 540;
 Best Local Similarity 57.9%; Pred. No. 4e-79;
 Matches 331; Conservative 64; Mismatches 95; Indels 82; Gaps 12;

587 CAHYKPPFCVAPSGVKGKDLSTYMPKFPDEBACQPCPCINCTHSCVDLDDKCPAEQ 646
 3 CAHFIDGPHCVKACPAVGLBNDTL-VKXADANAVQLCHPNCTRCCKPGLBGC- 58
 647 RASPLTISIAVY-GILVVVVLGVGFGILIKRQOKIKXTMRLLOETELVEPLTPSGA 705
 59 NSGKTPSIAGVGVGLCLVAVGIGLGLYLRH-HIVRKQTLRLLOERELVEPLTPSGE 117
 706 MPQAOQRILKETELRKVKVLGSGAGCTVYKGIWIPGENVKIPVALKVLRENTSPKANK 765
 118 APQOAHRLIKETEFKVKVKGAGFTYKGLMIPGENVKIPVALKELRENTSPKANK 177
 766 EILDEAYVAVGSPVYSRLGLCLTSTVOLVQLMPYGLLDHVRNRRGLSODLLNW 825
 178 EILDEAYVAVSVNPHVCRLLGICLSTVOLITQLMPYGLLDYIRHKONIGSYLLNW 237
 826 CMQIAGKSTLEDRVLVHRDLAARNVLYKSPNHVKTIDFGIARLLDDETYHADGKVP 885
 238 CVOIAGKMYLERRLVHRDLAARNVLYKTPQHKITIDFGIARLLDDETYHADGKVP 297
 886 IKMALESILARRFTQSDVWSYGVTWELMTGAKYDGIPIAREIDLEKGRLLPOPP 945
 298 IKMALESILARRFTQSDVWSYGVTWELMTGSKYDGIPIAREIDLEKGRLLPOPP 357
 946 ICTIDVYMIWVKCMWIDSECRPRELVSEFSRNARDPQRFVYIQ-NEDGAPASPLDSTF 1004
 358 ICTIDVYMIWVKCMWIDSECRPRELVSEFSRNARDPQRFVYIQ-NEDGAPASPLDSTF 417
 1005 YRSLEDDDMGDLVDAEYLVPOGGFPCDPAPGAGMVHRRHSSSTRSGGDLTLGLE 1064
 418 YRLMEEBDMEDIVDAEYLVPHQGF-----NSPST----- 449

